

DTIC FILE COPY

2

AFOSR-TR- 89-0627

**COMPUTATIONAL METHODS FOR NONLINEAR DYNAMIC
PROBLEMS IN SOLID AND STRUCTURAL MECHANICS:**

**Progress in the Theory and Modeling of Friction
and in the Control of Dynamical Systems with Frictional Forces**

AD-A208 763

FINAL TECHNICAL REPORT

March 31, 1989

AFOSR Contract F49620-86-C-0051

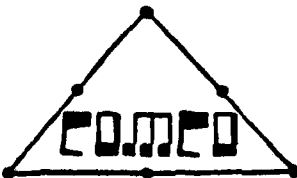
**Air Force Office of Scientific Research
Bolling Air Force Base, DC 20332**

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR)
NOTICE OF TRANSMISSION
THIS REPORT IS TRANSMITTED BY AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFOSR) AND IS
APPROVED FOR PUBLIC RELEASE AND FOR ACR 150-12.
DISTRIBUTION IS UNLIMITED.
MATTHEW J. KERPER
Chief, Technical Information Division

DTIC
ELECTE
MAY 16 1989
S D

TR-89-05

Approved for public release;
distribution unlimited.



THE COMPUTATIONAL MECHANICS CO., INC.
3701 North Lamar, Suite 201
Austin, Texas 78705
(512) 467-0618

1 2 0 5 1 5 1 2 2

**COMPUTATIONAL METHODS FOR NONLINEAR DYNAMIC
PROBLEMS IN SOLID AND STRUCTURAL MECHANICS:**
Progress in the Theory and Modeling of Friction
and in the Control of Dynamical Systems with Frictional Forces

FINAL TECHNICAL REPORT
March 31, 1989

AFOSR Contract F49620-86-C-0051

Air Force Office of Scientific Research
Bolling Air Force Base, DC 20332

TR-89-05



Accession For	
NTIS CRA&I	<input checked="checked" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	



THE COMPUTATIONAL MECHANICS CO., INC.
3701 North Lamar, Suite 201
Austin, Texas 78705
(512) 467-0618

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION Unclassified		1b. RESTRICTIVE MARKINGS None	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release, distribution unlimited	
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE			
4. PERFORMING ORGANIZATION REPORT NUMBER(S) TR-89-05		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR-89-0627	
6a. NAME OF PERFORMING ORGANIZATION Computational Mechanics Co, Inc	6b. OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR	
6c. ADDRESS (City, State, and ZIP Code) 3701 North Lamar Blvd, Suite 201 Austin, TX 78705		7b. ADDRESS (City, State, and ZIP Code) same as 8c.	
8a. NAME OF FUNDING / SPONSORING ORGANIZATION USAF, AFSC Air Force Office of Scientific Rsch	8b. OFFICE SYMBOL (if applicable) NA	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-86-0051	
8c. ADDRESS (City, State, and ZIP Code) Bolling Air Force Base DC 20332		10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 61024 PROJECT NO. 2302 TASK NO. B1 WORK UNIT ACCESSION NO.	
11. TITLE (Include Security Classification) "Computational Methods for Nonlinear Dynamic Problems in Solid and Structural Mechanics: Progress in the Theory and Modeling of Friction and in the Control of Dynamical Systems with Frictional Forces" (U)			
12. PERSONAL AUTHOR(S) J.T. Oden/ W.W Tvorzydlo/ J.A.C. Martins			
13a. TYPE OF REPORT Final Technical	13b. TIME COVERED FROM 11-1-88 TO 2-28-89	14. DATE OF REPORT (Year, Month, Day) March 31, 1989	15. PAGE COUNT 404
16. SUPPLEMENTARY NOTATION			
17. COSATI CODES FIELD GROUP SUB-GROUP		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) friction, damping, structural dynamics, elastohydrodynamic lubrication, (Mg)	
19. ABSTRACT (Continue on reverse if necessary and identify by block number) <p>This final report summarizes work on a three-year project devoted to dynamic friction and on control of systems with frictional forces. Detailed theoretical and numerical studies of static and dynamic friction and various phenomena of dynamic friction (stick-slip motion, self-excited oscillations, frictional damping, etc.) are described. Included in the report is also a study of methods of statistical homogenization, devised to develop new micromechanics-based models of contact and friction. The question of optimal control with distributed dynamical systems of viscoelastic bodies with frictional forces is also described. A final component of this study focuses on the mathematical and numerical analysis of elastohydrodynamic lubrication problems and is included in the final chapter.</p> <p>The studies presented in this report provide both a theoretical and a numerical basis for modeling contact and friction. These models are shown to produce results in very good qualitative and quantitative agreement with experimental results.</p>			
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT. <input checked="" type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Anthony K. Amos		22b. TELEPHONE (Include Area Code) (202) 767-4937	22c. OFFICE SYMBOL NA

TABLE OF CONTENTS

Chapter 1	Final Report on Modeling and Control of Friction	1
1.	Objective and Summary of Research	1
2.	Publications and Presentations	4
3.	Other remarks	5
Chapter 2	A Study of Static and Kinetic Friction	6
1.	Introduction	6
2.	Static and Kinetic Friction Stick-Slip Motion	8
2.1	Historical Background	8
2.2	Time Dependence or Rate Dependence of the Coefficient of Static Friction	13
2.3	The Steady-State Coefficient of Kinetic Friction	16
2.4	The Coefficient of Kinetic Friction During the Slip Phase of Stick-Slip Motion	17
2.5	Memory-Dependent Friction	17
2.6	The Importance of the Normal Degree of Freedom in Sliding Friction	19
	(I) Asymmetry of Normal Contact Oscillations	20
	(II) High-Frequency Stick-Slip Motions	20
3.	A Model of Interface Response	24
4.	Normal Oscillations and Stick-Slip Motion with a Sliding Friction System	29
4.1	Governing Equations	29
4.2	Numerical Results	32
4.3	Discussion	32
5.	Low Speed Dynamic Friction Phenomena with a Three-Degree- of-Freedom Rigid Body Model	43
5.1	Governing Equations	43
5.2	Steady-Sliding Equilibrium and Linear Stability Analysis	47
5.3	Low-Frequency Stick-Slip Motion and Apparent Reductions of Kinetic Friction	62
5.4	Apparent Reductions of Static Friction Due to Normal Perturbations	95

6. Critical Sliding of a Rigid Block on a Fixed Horizontal Plane	103
6.1 Governing Equations	103
6.2 Numerical Results	106
6.3 Discussion	118
7. Conclusions	121
Chapter 3 Numerical Study of Frictional Behavior of Sliding Systems	123
1. Introduction	123
2. Remarks on the Nature of Dynamic Friction	125
2.1 Dynamic Instability of the System	125
2.2 Rate and Time Dependence of the Static Coefficient of Friction	127
2.3 Dissipation of Energy During Normal Oscillations	128
2.4 Normal Jumps of a Slider	129
2.5 Load Dependence of the Coefficient of Friction	130
2.6 Dynamic Interlocking of Imperfections	130
3. General Presentation and Formulation of the Boundary-Value Problem	132
3.1 Equations for Body A	133
3.2 Presentation of the Moving Body B by a Frame F	135
3.2.1 Definition and Assumptions	136
3.2.2 Conservation Laws in Frame Description	139
3.2.3 Boundary Conditions for Body B	140
3.3 Model of the Interface	140
3.3.1 Definition and Assumptions	140
3.3.2 Kinematics of the Interface	142
3.3.3 Constitutive Relations for the Interface	143
4. Variational Formulation of a Boundary-Value Problem	145
5. Discretization of the Boundary-Value Problem	148
5.1 Finite Element Discretization	148
5.2 Incremental Equations of Motion	150
5.3 Rigid Body Model	150
5.4 Remarks on Methods of Solving the Equations of Motion	154
6. General Presentation of Numerical Examples	155

7. Summary of Experimental Results	156
8. Analysis of Stability of Frictional Sliding	159
8.1 Basic Computational Model	159
8.2 Estimation of Stability via Eigenvalue Analysis	161
8.2.1 The Basic Algorithm	161
8.2.2 Influence of Normal Stiffness on the Stability of Frictional Sliding	162
8.2.3 Influence of Other Parameters on the Frictional Stability	164
8.3 Transient Analysis of Self-Excited Oscillations	168
9. Introductory Analysis of the Kinetic Coefficient	172
10. Influence of the Angle of Attack on the Stability of Frictional Sliding	181
11. Conclusions	183
Chapter 4 New Models of Contact and Friction of Rough Surfaces	186
1. Introduction	186
2. Surface Topography Models	191
2.1 Random Microtopography Model	191
2.2 Numerical Calculation of Macrocontact Expectations	198
3. Adhesion	199
4. Contact Mechanics of a Viscoplastic Asperity	203
4.1 Variational Formulation	208
4.2 Numerical Analysis	210
5. Sliding Resistance	211
6. Future Work	213
Chapter 5 On the Controllability of Distributed Viscoelastic Systems With Frictional Forces	214
1. Introduction	214
2. Preliminaries	215
3. A Variational Formulation	217
4. Regularized Problems	223
5. Optimality Conditions for Regularized Problems	227
6. Approximation	234

Chapter 6 Qualitative and Numerical Analysis of Elastohydrodynamic

Lubrication Problems	238
1. Introduction	238
2. Fundamental Equations	241
2.1 Model	241
2.2 Simplifying Assumptions	343
2.3 Reynolds Equation	244
2.4 Elasticity	247
2.5 Boundary Conditions. Cavitation-Free Boundary	250
2.6 Summary	251
3. Variational Formulation — Penalty Method	253
3.1 Preliminaries on Variational Inequalities	254
3.2 Variational Inequality — Penalty Method	257
3.3 Penalized Formulation	259
3.4 Treatment of Free Boundary	260
4. Qualitative Analysis	261
4.1 Pseudomonotone Operators	261
4.2 Properties of Operator A	262
4.3 Existence Theorem	274
4.4 Brief Comments on Uniqueness	275
4.5 Convergence of the Penalty Method	275
4.6 A Regularity Theorem for Solutions of the Line Contact Case	279
5. Finite Element Solutions	287
5.1 Brief Survey on Finite Element Methods in Tribology	288
5.2 Finite Element Formulation	288
5.3 Iterative Schemes	290
6. Error Estimate	311
6.1 Approximations of Penalty Solutions in Finite- Dimensional Spaces	311
6.2 Approximation of Variational Inequalities in Finite-Dimensional Spaces	315
6.3 An A-Posteriori Error Estimate	318

6.4	Numerical Experiments on Error Estimates For Finite Element Solutions	321
6.5	Application of Adaptive Methods	330
7.	Concluding Remarks	365
7.1	Summary	365
7.2	Recommendations for Future Work	367
References	372

CHAPTER 1

FINAL REPORT ON MODELING AND CONTROL OF FRICTION

1 Objective and Summary of Research

The broad goals of this project, sponsored under Air Force Contract F49620-86-C-0051 and initiated three years ago, were to develop new mathematical models of friction that were consistent with experimental data and could be used in the calculation of the dynamics and control of structures subjected to contact and frictional forces. This general line of research was inspired by our work in earlier phases of this same project, when studies at COMCO revealed that a simple, normal compliance law for rough interfaces could be used that would not only produce a tractable mathematical theory, but also produce results that explained a long list of experimental results and that resolved several paradoxes that appeared to exist in experiments on dynamical friction, particularly long-held misconceptions on the nature and effects of the kinetic coefficients of friction. In the short time since their appearance, these friction laws have already had an impact on calculations in nonlinear structural dynamics, and have been incorporated in a number of codes to model dynamic friction effects.

These previous studies had also clearly indicated that the progress in understanding and realistic modeling of contact and friction would require combined effort toward developing new models of contact and friction of rough surfaces, developing dynamic models and methods of numerical analysis of mechanical systems with friction, experimental validation of theoretical and numerical predictions as well as continuous development of solid mathematical basis for modeling and control of systems with friction.

Thus, the scope of work in this project was broad, covering virtually all aspects of friction modeling from a detailed study of the micro-mechanics of surfaces and their contribution to frictional phenomena, a critical evaluation of a large volume of experimental and empirical work, the translation of these micro-mechanical observations into a workable mathematical theory, the use of such theories to characterize the constitution of interfaces prevalent in structural dynamics problems, the thorough numerical study of static and kinetic friction and phenomena of dynamic friction, the study of the mathematical properties of the formulation with particular emphasis on a mathematical tractability of the problem, well-posedness (existence and uniqueness of solutions and dynamical stability of the systems), and work on control theory (do controls exist for such systems with frictional forces?).

Of particular interest were classical frictional effects on structural response, such as stick-slip motion, self-excited oscillations, frictional damping, and damage of contact surfaces brought on by frictional forces and cyclic loading. Since, in practice, the effect of lubricants are also of interest, one important component of the present study also focuses on a mathematical and numerical analysis of the highly nonlinear Reynolds-Hertz elastohydrodynamics lubrication problem for non-Newtonian lubricants.

The particular problems resolved during three years of work on this project are:

1. A detailed study of the implications of the friction model on mechanisms responsible for apparent static and kinetic friction;
2. A numerical modeling of frictional behavior, in particular self-excited oscillations, of real sliding systems and quantitative comparison with experimental evidence;
3. A continued development of new methods of homogenization, including a statistical method of randomly distributed asperities which captures the principal micro-mechanical mechanisms responsible for friction damping and adhesion, and particularly those mechanisms which validate the macro-models developed during the first year of the effort;
4. Use of the new mathematical models of friction in studies of optimal control of elastic and viscoelastic structures;
5. Mathematical and numerical analysis of elastohydrodynamics problems for non-Newtonian lubricants.

To resolve the first problem, a finite element code for studying nonlinear response of structures with friction effects was developed. The macro-friction model developed by Oden and Martins earlier in this effort was used to model frictional effects on dry elastic metallic surfaces. A very detailed study of the effects of compliant asperities on real and apparent coefficients of friction was conducted. Once again it was confirmed that for elastic interfaces, the apparent reduction in the coefficient of friction is due to a separation of the compliant contact surfaces, due to transverse momentum components incurred during the motion. A study of these effects has led to a lengthy paper on the subject, which will possibly be published in monograph form. An abridged version of this work is included as Chapter 2 of this report.

In Chapter 3, the generalized Oden-Martins model of friction is applied in modeling of various phenomena of dynamic friction, in particular self-excited oscillations and the kinetic coefficient of friction. The particular objective of this effort was to model both qualitatively and quantitatively the behavior of real experimental pin-on-disk apparatus. Towards this end a special theoretical formulation and numerical program was devised for the analysis of

two-body contact with long distance sliding (pin sliding on a disk). In the course of numerical analysis, the mechanism of self-excited oscillations was clearly identified and parameters affecting their occurrence were studied. A strikingly good agreement with existing experimental evidence was observed, which confirms applicability of the Oden-Martins model and general methodology presented in this report to the modeling of frictional behavior of real engineering systems.

Another important development in the general area of friction was the construction of the micro-asperity-based model for deformable interfaces. This has been shown to lead to precisely the same macro-model used in all of the numerical experiences. However, because these micro-models address the particular micro-mechanical mechanisms responsible for the deformation of asperities, it is now possible to extend these theories to include thermal effects, viscous effects, plasticity, and damage. While some additional studies in this area have been made (and two papers on this subject have been accepted for publication this year), attention very recently has focused on the introduction of damage effects in the asperity model. Inclusion of damage in the model should provide a reasonable simulation of friction on brittle material surfaces and also provide for modeling the degradation of frictional resistance due to damage of material interfaces.

Also during this contract reporting period, work on the control of dynamical systems with friction has been completed. The basic issue is this: for a deformable viscoelastic body subjected to frictional forces, is it possible to define the program of applied surface forces on the body that will produce a given motion? Stated in another way, when frictional forces are present, is it possible to control the motion of a deformable body by any reasonable program of applied forces? This, until recently, was an open problem in distributed control theory and in the theory of partial differential equations. The principal source of difficulty was a tractable friction model itself; with Coulomb's law the mathematical issues are virtually intractable. On the other hand, we have recently shown that with the introduction of the Oden-Martins' law, a complete and tractable mathematical theory of optimal control of structures with friction can be developed. An interesting aspect of the results is that the presence of some viscous damping in the structure itself, i.e., viscoelasticity, appears to be needed to ensure well-posedness of the optimization problem. A summary of this work is given in Chapter 5 of this report.

Finally, in Chapter 6 our work on theory of lubrication is summarized. The strongly nonlinear Reynolds-Hertz equations and the classical free-boundary problems of elastohydrodynamic lubrication were formulated in a variational inequality framework and then regularized with the penalty method. The existence of the solution to this problem is proved. A penalty method was introduced to control the free boundary of cavitation and relevant finite element codes were generated. The numerical results show that the penalty method works excellently in locating the free boundary. Further study confirmed the convergence of

the penalty method and the finite-dimensional approximations. Moreover, an *a priori* error estimate for finite element solutions was proposed and an adaptive refinement scheme was developed for generating optimal meshes.

In the next section, we list papers that have been published and presentations made during the period of this project.

2 Publications and Presentations

The following publications and presentations have been made during this report period:

2.1 Publications

1. J. T. Oden, "New Models of Friction for Nonlinear Elastodynamics Problems." *Engineering Science Reprints*, October, 1987.
2. J. T. Oden, "New Models of Friction for Nonlinear Elastodynamics Problems." *Journal de Mécanique Théorique et Appliquée*, Special issue Supplement, No.1, to Vol 7, pp. 47-54, 1988.
3. S. R. Wu and J. T. Oden, "A Note on Applications of Adaptive Finite Elements to Elastohydrodynamic Lubrication Problems." *Communications in Applied Numerical Methods*, Vol. 3, pp. 485-494, 1987.
4. S. R. Wu and J. T. Oden, "A Note on Some Applications of Adaptive Finite Elements to Elastohydrodynamic Lubrication Problems." *International Journal of Engineering Science*, Vol. 25, No. 6, pp. 681-653, 1987.
5. S. R. Wu and J. T. Oden, "Convergence and Error Estimates for Finite Element Solutions of Elastohydrodynamic Lubrication," *Computers and Mathematics with Applications*, Vol. 13, No. 7, pp. 583-593, 1987.
6. P. J. Rabier and J. T. Oden, "Solution to Signorini-Like Contact Problems Through Interface Models, Part I: Preliminaries and Formulations of a Variational Equality," *Nonlinear Analysis, Theory, Methods and Applications*, Vol. 11, No. 12, pp. 1325-1350, 1987.
7. P. J. Rabier and J. T. Oden, "Solution to Signorini-Like Contact Problems Through Interface Models, Part II: Existence and Uniqueness Theorems," *Nonlinear Analysis, Theory, Methods and Applications*, Vol. 12, No. 1, pp. 1-18, 1987.
8. L. W. White and J. T. Oden, "Dynamics and Control of Viscoelastic Solids with Contact and Friction Effects." *Journal of Nonlinear Analysis*, (to appear).

9. J. T. Oden, "Dynamic Friction Problems", in *Contact Problems in Elasticity*, SIAM Publications, pp. 451-470, 1988.
10. J. A. C. Martins, J. T. Oden, and F. M. F. Simões, "Static and Kinetic Friction," *Wear*, (in review).

2.2 Presentations

1. J. T. Oden, "Paper at AFOSR/ARO," **Conference on Nonlinear Dynamics, Stability and Dynamics of Structures and Mechanisms**, Virginia Polytechnic and State University, Blacksburg, Virginia, March 23-25, 1987.
2. J. T. Oden, "Progress on Models of Friction and Contact," **Seminar, Air Force Office of Scientific Research**, Bolling Air Force Base, Washington, DC, May 4, 1987.
3. J. T. Oden, "New Models of Friction for Nonlinear Elastodynamics Problems," **25th Meeting of the Society of Engineering Science**, University of Utah, Salt Lake City, October 21, 1987.

3 Other Remarks

During the report period, the principal investigator, Dr. J. T. Oden, was assisted by the following COMCO Engineers and Consultants: Dr. J. M. Bass, Dr. T. H. Miller, Dr. W. W. Tworzydło, Mr. Chris Berry, Dr. L. W. White and Dr. P. J. Rabier. Dr. White collaborated with Dr. Oden on the work on optimal control theory discussed in Chapter 3, while Dr. Rabier has collaborated with Oden on several mathematical aspects under study, including the one listed in publication 6, and more recently, on modeling of damage models for metallic interfaces.

CHAPTER 2

A STUDY OF STATIC AND KINETIC FRICTION

1 Introduction

Most experimental studies on friction and wear are concerned with finding qualitative and/or quantitative relationships between friction and wear data and various governing parameters and conditions, for example: the properties of bulk and surface layer materials, the roughness of the surfaces in contact, the stress levels, the sliding speed, the temperature, the environment, the properties of the lubricants and the lubrication conditions. The correlations found in those experimental works are usually assumed to be intrinsic characteristics of the interfaces tested for the ranges of parameters and conditions considered. Difficulties on the reproducibility of friction data with different experimental apparati under otherwise similar conditions, and dependence of the results of wear tests on the dynamic properties of the equipment have been occasionally mentioned in the literature ([19], [48], [54], [69]), but only in the last few years has an increasing attention been paid to the experimental study of these effects ([2], [3], [20], [21], [22], [27], [62]).

It is the objective of this work to present and discuss various numerical studies on low speed frictional sliding phenomena which strongly suggest that classical interpretations for these phenomena should be critically re-examined and the role played by the dynamic properties of the experimental apparati should be carefully taken into account. In most of the discussions we shall restrict ourselves to dry frictional sliding of metallic bodies and particular emphasis will be given to the importance of normal (and rotational) degrees-of-freedom in dynamic friction phenomena.

This work is essentially a continuation of our previous paper (Oden and Martins [56]) where a thorough review of various aspects of frictional sliding phenomena and interface behavior was presented, a simple constitutive model for the interface was proposed and some preliminary numerical studies on small speed sliding phenomena were done.

Following this introduction, we review in Section 2 some aspects of low speed frictional sliding phenomena, particularly those related to distinctions between static and kinetic friction, the dependence of the latter on the sliding speed and the dependence of the static friction on the time of stationary contact. In the same section, special attention is given to the experimental observations and ideas of Tolstoi [74] and Budanov, Kudinov, and Tolstoi [18] on the fundamental role played by the normal degree-of-freedom in frictional sliding phenomena.

These observations and ideas have a marked influence on the approach followed here, particularly on the interface model ultimately presented and discussed in Section 3. As suggested by the results of the experiments of those Russian authors, we do not make any distinction between the coefficients of static and kinetic friction.

In Section 4, a simple two degree-of-freedom model is used to simulate stick-slip tangential oscillations induced by normal vibrations. The effect of the normal jumps occurring during the slip phase on the friction-sliding velocity loops is demonstrated through numerical simulations.

In Section 5, we continue the study initiated by Oden and Martins [56] on the dynamic behavior of a three degree-of-freedom rigid body slider subjected to its own weight and restrained by a tangential spring and dash-pot arm. A nondimensional form of the governing system of equations and inclusions is derived and the governing nondimensional parameters are identified. Linear stability analyses of the steady-sliding equilibrium positions are presented and the regions of dynamic instability in the parameter space are identified. Consequences of the dynamic instability of the steady sliding are then studied. For small driving speeds, small tangential stiffness and small tangential damping, low-frequency stick-slip motions are observed in the numerical simulations. The corresponding plots of the stick-slip amplitude versus driving speed and frequency versus driving speed are presented and qualitative comparisons with experimental observations are provided. For sufficiently large driving velocity or for sufficiently large tangential stiffness or damping, "apparently smooth" sliding motions at apparent values of the coefficient of kinetic friction smaller than the static one are numerically obtained. The corresponding plots of the variation of the apparent coefficient of kinetic friction with the average sliding speed are presented and discussed. In Section 5 we also study the effect of normal perturbations along the stick phase of stick-slip motions on the measurable value of the coefficient of static friction. The role of the rate of increase of the tangential force on the computed results is also discussed.

In Section 6, we essentially follow an idea of Euler [33] for an experiment on the distinction

between static and kinetic friction: we consider a body on a fixed horizontal plane acted upon by a horizontal force equal to the critical force of static friction. Contrary to classical interpretations, we show that no distinction between the coefficients of static and kinetic friction is required in order for the body to have an (at least initially) accelerated motion.

Finally, in Section 7, the major conclusions of this work are summarized.

2 Static and Kinetic Friction. Stick-Slip Motion

2.1 Historical Background

Distinctions between coefficients of static and kinetic friction have been mentioned in the friction literature for centuries, at least since the work of Euler [33]. Indeed, Euler, whose major contribution to friction science was his elegant application of the general laws of mechanics and the available calculus machinery to the solution of various friction problems, also developed a mechanical model to explain the origins of frictional resistance. In so doing, he arrived at the conclusion that friction during sliding motion should be smaller (one-half, he suggested ...) than friction at the onset of sliding. Euler even proposed an experiment to check the truth of his conjecture and worked out the details of the solution of the corresponding mathematical problem.

The experiment proposed by Euler involved the sliding of a body down an inclined plane at slopes close to the critical slope at which sliding initiates. He expected the following “paradox” to be observed: it would not be possible to give to the inclined plane inclinations such that the descent would be as slow as desired — either the body would not slide at all or it would slide very fast. This, of course, would mean that, as soon as sliding initiates, a drop in friction force occurs, the difference between static and kinetic friction forces being responsible for the acceleration of the body down the inclined plane.

The distinction between static and kinetic friction was also a major topic of Coulomb’s [24] detailed experimental study. Coulomb’s work is, in addition, the first major reference dealing with the *increase of the coefficient of static friction with increasing times of repose* (stationary contact before the initiation of sliding). Indeed, for certain combinations of materials and surface conditions, Coulomb observed distinctions between static and kinetic friction, a dependence of the kinetic friction on the sliding velocity and a dependence of the static friction on the time of repose. However, *for dry metal-to-metal interfaces all those distinctions or variations were absent or negligible*.

Shortly after Coulomb’s work, Vince [78] also observed that, for a variety of hard ma-

terials, the coefficient of kinetic friction was independent of sliding speeds. Through the nineteenth century, various authors confirmed the observations of Coulomb for dry metallic interfaces with regard to the coefficients of static and kinetic friction (Rennie [61], Morin [55], Hirn [38], and Jenkin and Ewing [44]).

Other authors, however, had different views: Kimball [49] and Conti [23] proposed, on the basis of their experiments with various dry or lubricated surfaces, that, in general, the coefficient of kinetic friction would be small and increasing with sliding velocity at low velocities. Then, at some velocity (dependent on the materials and the normal pressure), it would achieve a maximum value after which it would decrease with the increase of speed. But also from Conti's experiments it was clear that such variations were smaller for the case of dry interfaces.

As noted by Kragelskii [50], the early experiments on kinetic friction were done at relatively small sliding speeds: Coulomb did not exceed 2.5 m/s; Rennie 2.56 m/s, Morin 4 m/s and Jenkin and Ewing 0.003 m/s. The application of dry friction in the brakes of railway carriages prompted the study of frictional sliding at higher velocities. For sliding speeds in the range of approximately 1–25 m/s Poirée [57], Bochet [10], and Galton [34] observed decreases of the coefficient of friction with the increase of sliding velocity, consistent with the results of Kimball and Conti. For sufficiently high sliding speeds it is thus clear that decreases of kinetic friction with increasing speeds may indeed occur in the dry sliding of metallic bodies. Kragelskii [50, 51] provides various practical formulae for this situation and explains it as being the result of *material softening due to high temperatures generated on the contact neighborhood*.

However, for the small velocity range that we are mostly interested in, the situation is not so clear. Conflicting results obtained with various lubrication conditions and the absence of a clear understanding on the distinction between dry and lubricated sliding added much to the confusion established during the second half of the past century on the velocity dependence of the coefficient of friction. Unfortunately, the fundamental advances on the theory and applications of lubrication by the turn of the century (1880's to 1910 — see Dowson [29]) apparently were not accompanied by corresponding advances on the knowledge of dry friction at small velocities.

Well into the twentieth century Bowden and Leben [11] and Bowden and Tabor [12], while studying the nature of kinetic friction between dry metallic surfaces, observed that the friction force was not constant during sliding:

“Experiments with dissimilar metals sliding together show that as a general

rule the frictional force is indeed not constant but undergoes most rapid and violent fluctuations. The sliding process is not a continuous one; the motion proceeds by jerks. The metallic surfaces "stick" together until, as a result of the gradually increasing pull, there is a sudden break with a consequent very rapid "slip". The surfaces stick again and the process is repeated indefinitely. When the surfaces are of the same metal, the behavior is somewhat different. Large fluctuations in the friction still occur but they are comparatively slow and very irregular. The average value of the frictional force is considerably higher than that found for dissimilar metals and a well-marked and characteristic track is formed during the sliding" (Bowden and Leben [11]).

The stick-slip motions described in the first part of the above quotation have a sawtooth wave form (see Figs. 2.1. and 2.2) and are the typical friction-induced oscillations observed at small sliding speeds.

The need to eliminate or attenuate intermittent ("stick-slip") motions similar to those described in the first part of the above quotation in various practical applications, has originated the publication of a large number of studies on the subject during the past fifty years. These studies have provided most of the recent information on dry sliding friction at small speeds. The papers by Bell and Burdekin [5], Antoniou, Cameron and Gentle [1], Richardson and Nolle [64], and Oden and Martins [56], provide surveys of related aspects of the friction literature and additional references.

Soon it was realized that the stick-slip motion described by Bowden and Leben was a relaxation oscillation which was influenced not only by the nature of the surfaces in contact, but also, in a fundamental manner, by the dynamic properties (stiffness, inertia, damping, ...) of the experimental apparatus.

In experiments carried out, using either specially designed apparatus or slightly modified machine tool tables and slideways, it has been observed that the amplitude of the stick-slip motion decreases when:

1. the driving velocity \dot{U}_T increases ([15, 47, 60])
2. the damping coefficient C increases ([15])
3. the spring stiffness K increases ([60])
4. the mass M of the slider decreases ([47]).

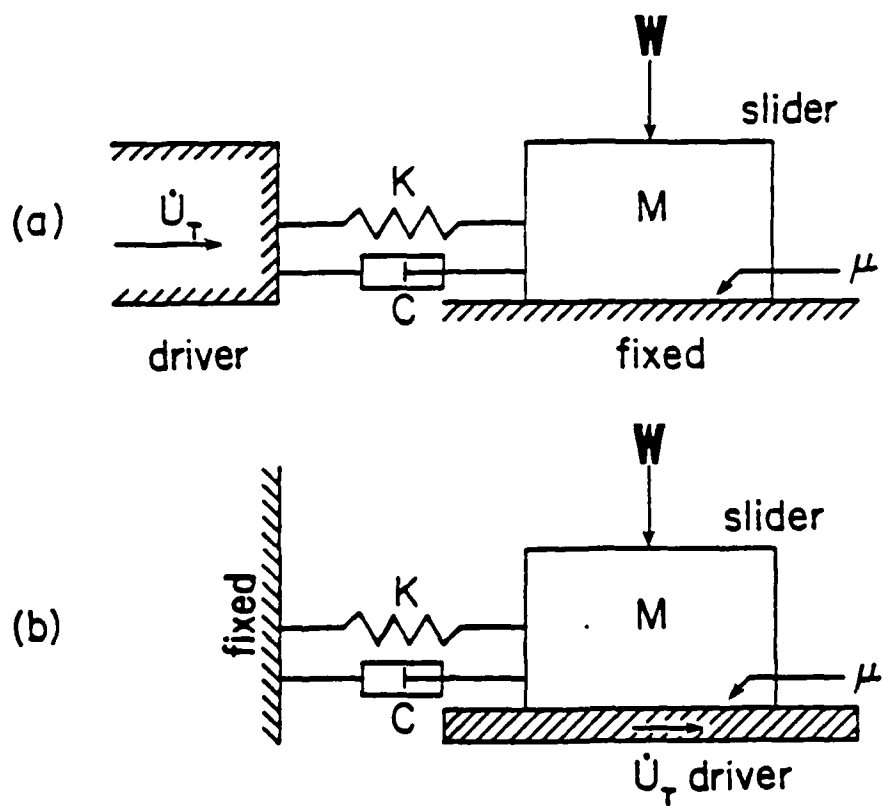


Figure 2.1: Models of two (equivalent) sliding systems which may have stick-slip oscillations. K = linear stiffness; C = linear damping; M = mass of the slider; W = normal load; \dot{U}_T = driving velocity; μ = friction coefficient.

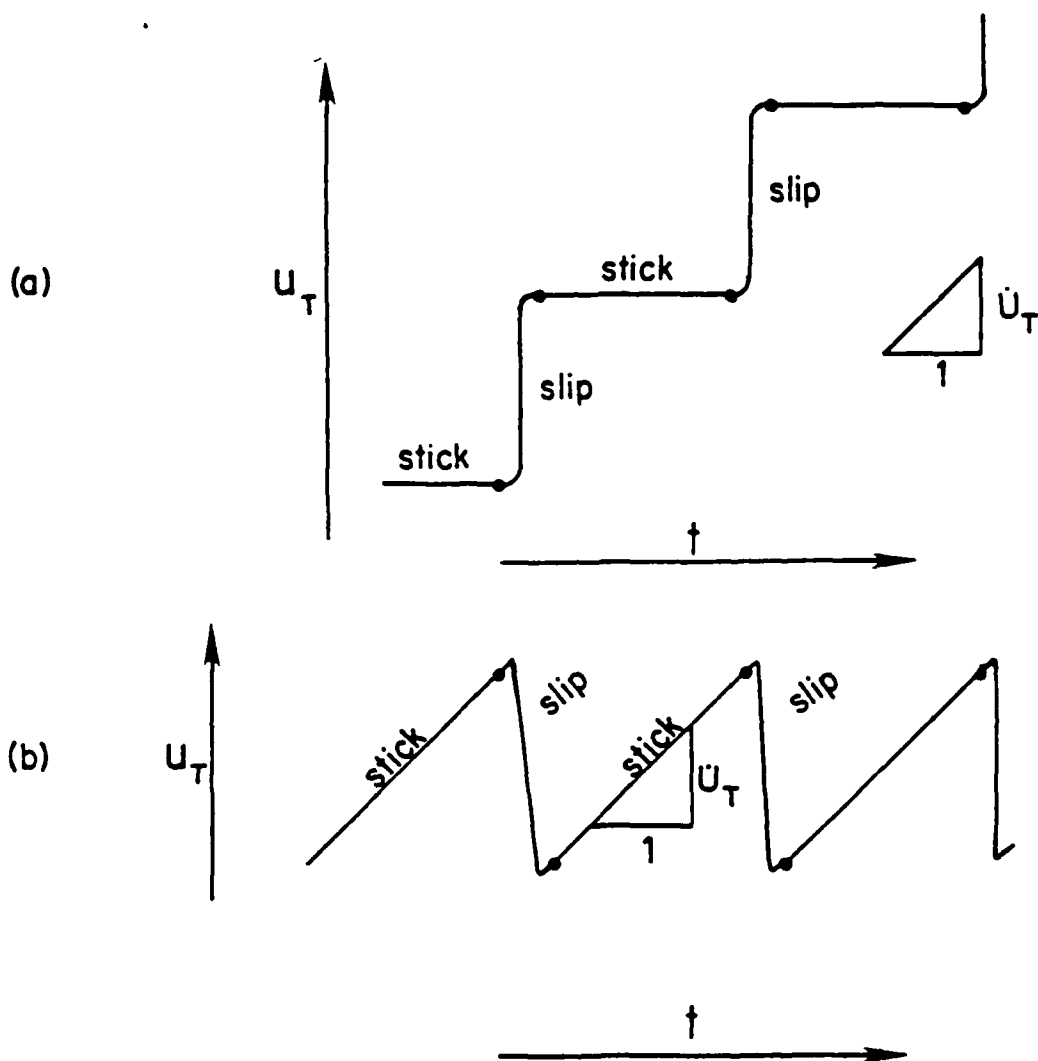


Figure 2.2: Typical traces of stick-slip motion for systems (a) and (b) in Figure 2.1.
 u_T = displacement of the slider; t = time; \dot{u}_T = driving velocity.

It has also been observed that the frequency of the stick-slip motion increases with the increase of the driving velocity and that the maximum value of this frequency approaches the undamped natural frequency of the system ([5], [28]) although in some cases the oscillation stops at a level well below that natural frequency.

Kaidanovskii and Khaikin [46] and Blok [8] pointed out that such oscillations might occur if the friction force decreased when the sliding velocity increases, a condition that may lead to an overall negative damping in the sliding system. *The decrease of the coefficient of friction with the increase of sliding velocity* according to some continuous or discontinuous law has been thus one of the most common assumptions in the studies of stick-slip motion. The other major assumption used in the literature is the *increase of the coefficient of static friction with the time of stationary contact* (Ishlinskii and Kragelskii [42]).

At this point a question arises: *if, as noted previously, most early researchers of frictional sliding with dry metallic bodies could not find significant dependence of the kinetic friction on the sliding speed at small speeds, or significant dependence of the static friction on the time of stationary contact, then what are the additional experimental results that support the use of such assumptions in the analyses of stick-slip motions?*

2.2 Time Dependence or Rate Dependence of the Coefficient of Static Friction

Consider a slider resting on a surface with no macroscopic sliding motion relative to the surface and the friction force Σ_T increasing at a constant rate $\dot{\phi} = \dot{\Sigma}_T/|\Sigma_n|$ until gross sliding occurs (Σ_n denotes the compressive normal contact force). Under these conditions, it can be observed that the value μ_s of $\phi = \Sigma_T/|\Sigma_n|$ at which the macroscopic sliding occurs, increases with the decrease of the rate ($\dot{\phi}$) of application of the tangential force (see Fig. 2.3). Observations of this kind can be done in the course of stick-slip oscillations: smaller driving velocities imply smaller rates of application of the tangential force and, consequently, the friction force at the onset of sliding becomes larger.

As seen in Fig. 2.3, smaller rates $\dot{\phi}$ correspond also to larger times of stationary contact (t_s). This led to the classical statement: the coefficient of static friction (μ_s) increases with the time of stationary contact (t_s). And this also led to a physical interpretation analogous to the one used to explain coefficients of static friction larger than coefficients of kinetic friction: *the strength of the contact junctions would increase with the time of stationary contact*. Indeed, according to the most widely accepted theories on the origin of friction (Bowden and Tabor [13, 14] and Rabinowicz [60]) the frictional resistance depends on the strength of

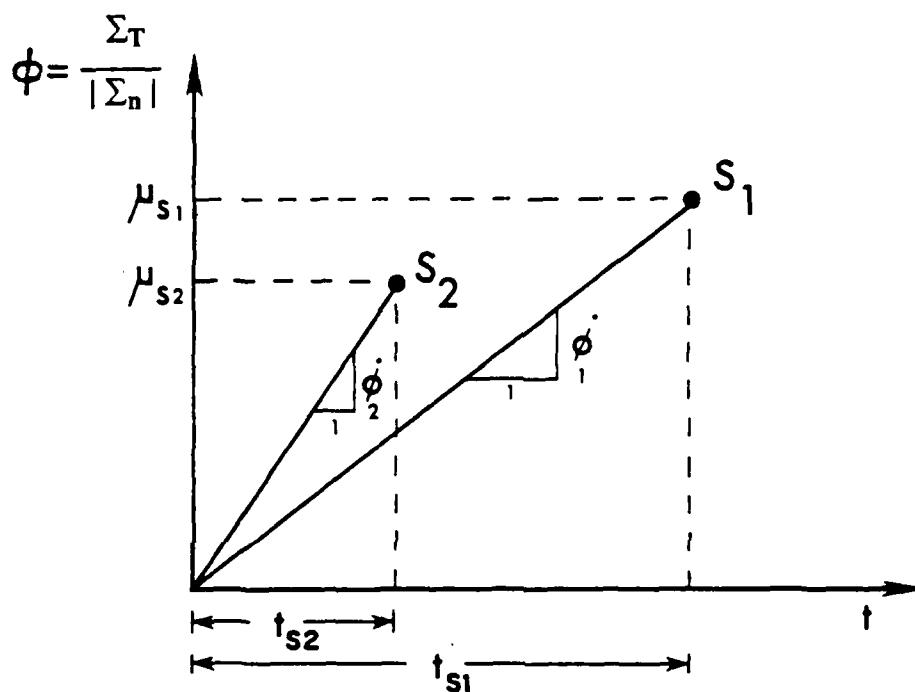


Figure 2.3: Rate dependence of the static coefficient of friction. S_1 and S_2 are the points at which gross sliding initiates; $\dot{\phi}_1$ and $\dot{\phi}_2$ are the rates of increase of the tangential force coefficients ϕ_1 and ϕ_2 ; t_{s1} and t_{s2} are the times of stationary contact; μ_{s1} and μ_{s2} are the static coefficients of friction.

the contact junctions formed between opposing asperities of the surfaces in contact; if this strength increases with the time of contact, then larger times of stationary contact imply larger coefficients of static friction; furthermore, during sliding, the time of contact between opposing asperities is much smaller than before the initiation of sliding, which would explain (Rabinowicz [60]) the occurrence of coefficients of kinetic friction smaller than the coefficient of static friction. Expressions proposed by several authors for this "time dependence of the coefficient of static friction" can be found in a survey paper by Richardson and Nolle [64] and have been used in the analysis of stick-slip motions. See also [16] and [47] for specific mechanisms proposed to explain the contact strengthening.

However, the experimental work of several authors suggests that these interpretations were not correct.

Simkins [68] carried out experiments to observe the micro-displacements of a slider before gross-sliding. He found that higher rates of loading inevitably led to macroscopic sliding at lower force levels. However, in other experiments designed to assess the influence of the time of stationary contact on the value of the static coefficient of friction, he could not find any correlation between the time of stationary contact and the value obtained for that coefficient.

Johannes, Green and Brockley [45] (working with lubricated surfaces) and Richardson and Nolle [64] (with "quite dry but not grease free" surfaces) carried out experiments in such a manner that they could vary independently the rate of application of the tangential force and the time of stationary contact. In those circumstances they found that *the governing variable was the rate of increase of the tangential force and not the time of stationary contact.*

The dependence of μ_s on $\dot{\phi}$ obtained by Richardson and Nolle is such that, *for sufficiently small load rates, the coefficient of static friction is constant and equal to a value which is the usually quoted coefficient of static friction. For sufficiently large loading rates the coefficient of static friction tends to be constant and equal to a value which is usually interpreted as the coefficient of kinetic friction, although comparative measurements have rarely been made. For intermediate load rates, the coefficient of static friction decreases with increasing load rates, providing a smooth transition between the above limit cases.*

As a consequence of their observations, Richardson and Nolle [64] suggest that empirical expressions of μ_s as a function of t_s should be recast as μ_s as a function of $\dot{\phi}$. Although such a program appears feasible, the implications of those experimental findings are more profound than that solution suggests: if the coefficient of static friction is not affected by the time of stationary contact, all the classical interpretations in terms of an increase of the strength of the junctions with time will no longer be valid. What is then the mechanism

responsible for the "rate dependence" of the static friction? Although it has been suggested (Bhushan [7]) that the strain rate dependence of the metal strengths should be taken into account in this context, to our knowledge no detailed explanation has been advanced and the whole subject appears to be far from settled (in this respect see also Tudor and Bo [77]).

2.3 The Steady-State Coefficient of Kinetic Friction

For hard on soft metal combinations (steel on indium and steel on lead) coefficients of friction increasing with sliding velocity in the ranges 10^{-10} to 10^{-4} cm/s and 10^{-10} to 10^{-8} cm/s , respectively, were obtained by Burwell and Rabinowicz [19]. For the same metal combinations the same authors found that the coefficient of friction decreases with the increase of sliding velocity for velocities greater than about 10^{-3} cm/s .

Friction-velocity plots of this type (initially increasing and then decreasing) are considered by Kragelskii [50] to represent the typical behavior of the steady-sliding coefficient of friction. The initially increasing portion is attributed by Rabinowicz [60], Kragelskii [50] and Tolstoi [74] to the creep deformation of the interface asperities. As already mentioned, the decreasing portion is explained by Kragelskii [50] and other authors as the result of thermal softening.

These views essentially agreed with the suggestion by Burwell and Rabinowicz [19] that, for hard metal pairs (aluminum, copper, steel, etc.) an initially increasing portion of the friction-velocity plots might also exist. The difficulty to provide experimental evidence of this at velocities of the order of 10^{-9} cm/s ($< 5 \times 10^{-10 \pm 1} \text{ cm/s}$ for steel pairs, according to Tolstoi [74]) is obviously extraordinary (see, however [74]). In any case, since those velocities are so small, we conclude that *for hard metal pairs, creep sliding, if possible at all, will have a negligible role on most common friction phenomena*. This essentially agrees with the view expressed by Bowden and Tabor [14] on the small part played by creep on the frictional behavior of ordinary engineering metals at room temperature.

On the other hand, despite the frequent allusions to coefficients of friction decreasing with sliding speed at small speeds, most of the experimental steady-sliding results available in the literature apply to lubricated surfaces. A reason for this, in addition to the obvious importance of the lubricated case, is the difficulty in obtaining, with dry metallic interfaces and most of the experimental apparatus, a smooth steady-sliding (without stick-slip oscillations) at small sliding speeds ([37]). Despite these difficulties, some steady-state decreasing friction velocity "characteristics" are reported in the works of Rabinowicz [60] (a small negative slope for titanium on titanium in the range of 10^{-7} to 100 cm/s , a larger slope for steel

on steel in the range of 10^{-3} to 100cm/s , Bell and Burdekin [5] (cast-iron on cast-iron for speeds smaller than 2.54 cm/s down to a nonspecified speed.) In another paper, Heyman, Rabinowicz, and Rightmire [37] concluded that the limited data obtained by them at that time suggested that, for most metals, the coefficient of friction was affected very little as the speed varied in the range 10^{-4} to 10^{-6}cm/s . Rabinowicz [60] states that for hard metal combinations decreasing μ/v_T curves are typically found but he emphasizes the smallness of the slopes of those curves when μ is plotted against the logarithm of the speed.

2.4 The Coefficient of Kinetic Friction During the Slip Phase of Stick-Slip Motions

Whatever the steady-state friction-velocity curve is or is believed to be, it soon became clear that during the slip portion of the stick-slip cycles the friction force would not follow the path predicted by such a curve. Instead, experimental results show that the friction force follows a loop — the friction force during the acceleration portion of the sliding is in general distinct from the friction force during the deceleration.

Unfortunately, the various experimental observations of these cycles are not conclusive: different material combinations and different experimental apparatus originate loops with distinct shapes and orientations ([1], [5], [39] and [81]) and, *even for the same materials and the same experimental apparatus, changes on the driving velocity or the dynamic properties of the apparatus affect radically the resulting loops* (see the experimental results of Bell and Burdekin [5] [Fig. 3, page 549]). This suggests that the friction-velocity plots obtained in the course of stick-slip motions are not an intrinsic property of the surfaces in contact—they are greatly affected by all the dynamic variables involved in each particular experimental set up.

2.5 Memory-Dependent Friction

In view of the complexity of friction phenomena at small speeds, new ideas were needed which might give a unified explanation for the apparently disparate experimental observations.

One such idea was advanced by Rabinowicz [59]: the friction force would be determined not only by the instantaneous sliding conditions but by the sliding history of a preceding critical distance (of the order of 10^{-3} cm for various metal surfaces). This critical distance concept, suggested by various experimental observations summarized in the same paper, could be used, according to its author, to give a unified explanation for different aspects of

sliding at low speeds.

Clockwise loops described by the coefficient of friction during the sliding portion of stick-slip experiments would be explained in the following manner: assume that, for steady-state sliding, the friction coefficient decreases with the increase in sliding velocity; then, if the instantaneous coefficient of friction is determined by the average velocity for the previous critical distance slid, it results that, when accelerating, the instantaneous coefficient of friction is larger than the steady-state coefficient corresponding to the same sliding velocity, the opposite occurring during the deceleration.

Increases of the coefficient of static friction (μ_s) with increasing times of stick (t_s) and decreases of the coefficient of kinetic friction (μ_k) with increasing sliding velocities (v_T) would be related by using the critical distance (d_c) concept:

$$|v_T| = d_c/t_s \Rightarrow \mu_s(t_s) = \mu_k(|v_T|).$$

Still, according to the same author, the transition from stick-slip motion to a smooth steady sliding would occur for a tangential stiffness sufficiently large that the corresponding stick-slip motions would have, under "plausible assumptions", amplitudes smaller than the critical distance d_c .

Although Rabinowicz [59] provided some promising comparisons between the predictions of his model and experimental results, the absence of a detailed analytical or numerical study on the behavior predicted by his model precludes a definitive conclusion on its validity. Furthermore, experimental results published after Rabinowicz's paper raise some new difficulties: How can one explain the counterclockwise loops of the friction coefficient observed by Antoniou, Cameron and Gentle [1]? How can one surmount the questions raised (recall Section 2.2) on the validity of the major physical basis for Rabinowicz's correlation between the $\mu_s(t_s)$ and $\mu_k(v_T)$ curves, i.e., the increase of the strength of the junctions with the time of contact?

More elaborate models which also take into account memory effects in frictional phenomena, have been advanced in recent years by Ruina, et. al. [35, 63, 65, 66, 84], and some promising results and simulations of geological fault slip phenomena based on these models have been presented (Tse and Rice [76]). A brief summary and discussion of this current field of research can be found in Oden and Martins [56].

2.6 The Importance of the Normal Degree of Freedom in Sliding Friction

Substantially different ideas were advanced mainly by some Russian authors: initially Kudinov [52] for lubricated contacts and later Tolstoi [74] for dry contacts. Tolstoi observed that the forward movements of a slider during stick-slip motion occur in strict synchronism with upward normal jumps. Observed decreases of friction during the sliding portions of the stick-slip motion might be thus the result of a decrease of the average normal contact force during the sliding and jumping, without the need to consider any reduction on the real coefficient of friction. More detailed observation of these upward jumps revealed that, while sliding, the body undergoes a normal oscillation, the frequency of which (of the order of 10^3 Hz) is consistent with the normal interface stiffness properties.

We note that several authors have also done observations analogous to these. In their early study, Bowden and Tabor [12] measured the interface electrical conductance during stick-slip motion and observed a marked fall of electrical conductance during the slip phase of the motion (Fig.9a, 9b, Plate 26, op. cit.). More detailed analysis also revealed that during those slips the conductance actually oscillated very rapidly with frequencies of the order 10^5 Hz (Fig. 10a, Plate 26, op. cit.). These changes in conductance are attributed by those authors to corresponding changes in the true area of contact and this indeed suggests the occurrence or normal oscillations of the type observed by Tolstoi and coworkers. Of course, for clearly distinct experimental apparatus, contact geometries and loads, the observed frequencies were also very different. Sharp decreases of electrical conductance during the slip phase of (lubricated) stick-slip motions can also be found in the work of Johannes, et al [45]. Direct measurements of the separation of unlubricated and lubricated surfaces during stick-slip motion were also done by Bo and Pavelescu [9] and by Tudor and Bo [77], respectively. Other experimental evidence, although less conclusive, of the influence of the normal degree of freedom on stick-slip motions was presented by Antoniou, Cameron and Gentle [1] and by Elder and Eiss [32].

Especially important was the experimental observation by Tolstoi that external damping of the normal free microvibrations of a slider could eliminate the decrease of friction force with the increase of sliding velocity. No quantitative distinction between static and kinetic friction could be observed under those circumstances. Furthermore, when normal damping was introduced, with no change on the driving velocity, during a run that showed stick-slip oscillations, it was observed that the oscillations ceased and that the value of the coefficient of friction for the subsequent smooth sliding was even greater than the maximum values

obtained at the end of the stick periods of the stick-slip motion. The responsibility of the freedom of normal displacement for both the falling friction force-sliding velocity relation and the stick-slip motion was corroborated in another manner: sufficiently heavy tangential damping alone could indeed suppress the stick-slip oscillations but it failed to affect the negative slope of the friction force-sliding velocity curve.

From these (and other) observations Tolstoi concludes that oscillations normal to the contact surface play a key role in both "static" and "kinetic" friction.

With respect to *static friction*, apparent reductions of the measurable static coefficient at any rate of application of the tangential force, would be the result of microseisms of amplitudes 0.1 to 10 μm commonly observed on the earth crust (Coulomb [25]).

With respect to *kinetic friction*, Tolstoi and coworkers have proposed the following mechanisms to explain apparent coefficients of kinetic friction lower than the static one:

(I) Asymmetry of normal contact oscillations [74]:

An increase in the speed of the slider increases upward components of the impulses exerted on the slider asperities as they collide with those of the underlying surface; this increases the amplitude of the normal natural vibrations of the slider which are governed by the contact stiffness and mass of the slider; due to the nonlinearity of the normal force-penetration relationship (see Fig. 2.4), the normal vibrations of the slider are highly asymmetric and, consequently, an increase of the amplitude of the oscillation decreases the mean level of penetration during sliding; hence, the average area of contact decreases and, as a result, the friction force also decreases. This mechanism provides thus an explanation for *apparent decreases of kinetic friction with increasing sliding speeds*. Tolstoi, Borisova and Grigorova [75] specifically propose the following "time averaged" law for kinetic friction:

$$|\Sigma_T^{aver}| = \mu_s |\Sigma_n(a^{aver})| \quad (2.1)$$

where Σ_T^{aver} denotes the time average of the friction force, μ_s denotes the coefficient of static friction, a^{aver} denotes the average penetrating approach and $-\Sigma_n(\cdot)$ denotes the (nonlinear) monotonically increasing function that characterizes the normal interface response (see Fig. 2.4).

(II) High-frequency stick-slip motions [18]

Oscillations in normal contact force induce similar oscillations on the maximum instantaneously available friction force so that during each cycle of normal oscillation the slider will

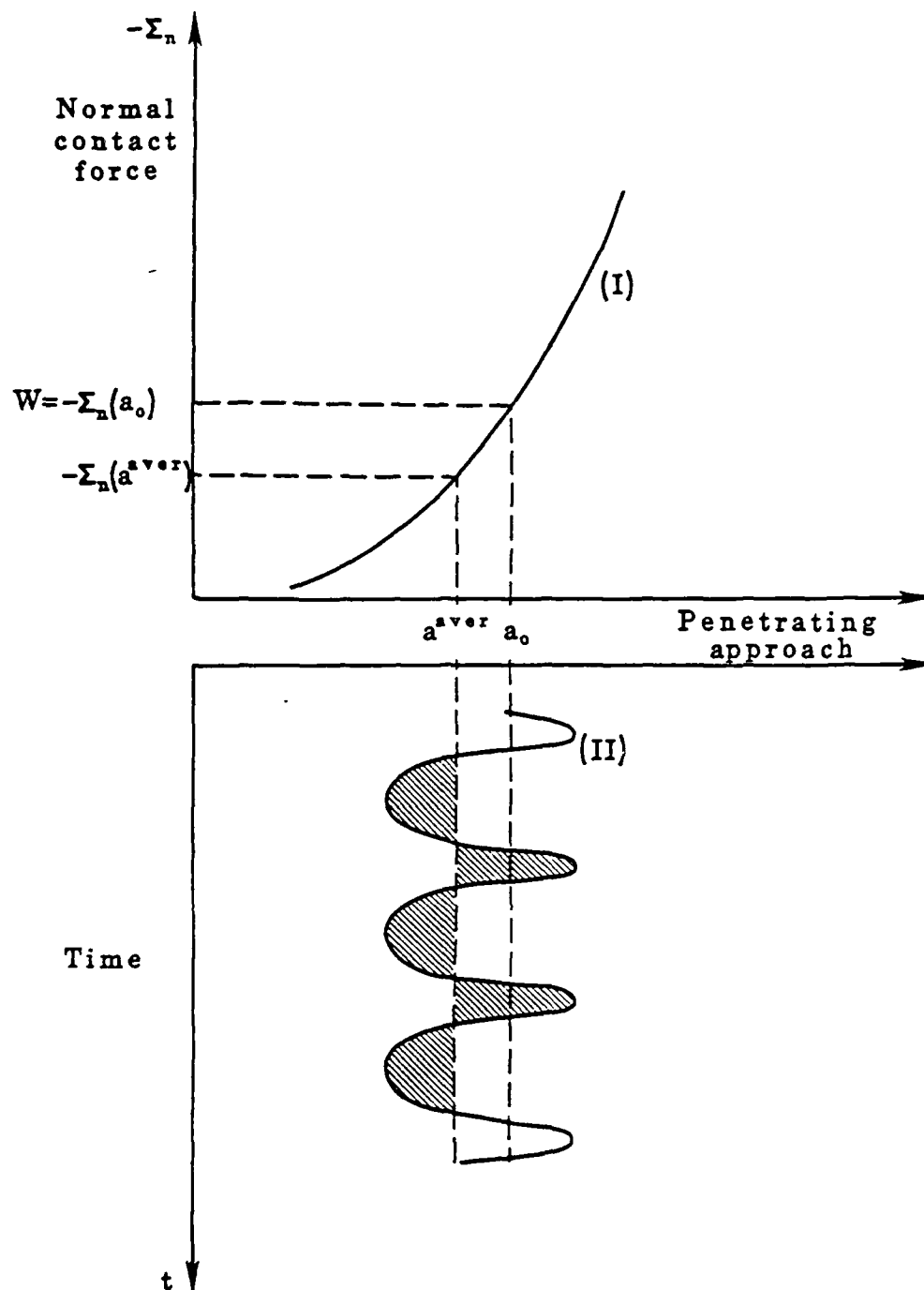


Figure 2.4: Static equilibrium penetrating approach (a_0) and average penetrating approach (a_{aver}) in the course of asymmetric normal oscillations. (I) curve characterizing the nonlinear elastic behavior of the interface in the normal direction (W is the applied normal load); (II) asymmetric normal free oscillation (shaded areas are equal.)

alternately stick and slide; since the frequency of the normal oscillation is high, its amplitude small and the average sliding velocity of the body also small, the motion of the body will be recorded as an apparently smooth sliding; furthermore, the ratio between friction and normal force for the short periods of sticking is smaller than the coefficient of static friction, so that, in average, an apparent coefficient of kinetic friction results which is smaller than the true static one; also, for larger average sliding speeds, the average time of stick of the successive stick-slip cycles will be smaller, so that *the average apparent coefficient of friction will increase with the average sliding velocity*.

Taking into account both mechanisms (asymmetry of normal oscillations and high frequency stick-slip motions), the resulting plot of "average apparent coefficient of kinetic friction" versus average sliding speeds would be, according to Budanov, Kudinov and Tolstoi [18], of the form depicted in Fig. 2.5.

We remark that the form of the curve coefficient of kinetic friction-sliding velocity depicted in Fig. 2.5 is consistent with what has been considered by Kragelskii [50, 51] the general behavior of the coefficient of kinetic friction (recall Section 2.3). However, owing to the nature of the mechanisms proposed by Tolstoi and coworkers to explain it, it results that: (i) the (apparent) coefficient of kinetic friction would be highly dependent on the dynamic properties of the experimental apparatus used to measure it; (ii) for hard metals, an increase of kinetic friction with sliding velocity might be observable for velocities that, although small, would not need to be so small as the validity of a creep sliding mechanism would require; (iii) decreases of kinetic friction with the increase in velocity might be observable for speeds not so large as a thermal softening effect would suggest; (iv) all of these *apparent* variations in kinetic friction could be reduced or eliminated by appropriately restraining normal degrees of freedom of the sliding system used.

It is clear from the above exposition that the ideas of Tolstoi aim at explaining in a unified and simple manner various aspects of frictional sliding at small speeds — apparent reductions of static and kinetic friction and stick-slip oscillations.

The experimental evidence pointing towards the importance of normal oscillations in frictional sliding, for a large range of experimental conditions, appears to be irrefutable (see our previous paper: Oden and Martins [56], for a summary of further experimental results by various authors).

A difficulty, however, arises when trying to formulate interface models consistent with Tolstoi's ideas: equation 2.1 postulates some average behavior in the course of sliding with asymmetric normal oscillations, but it is not a deterministic model of interface behavior.

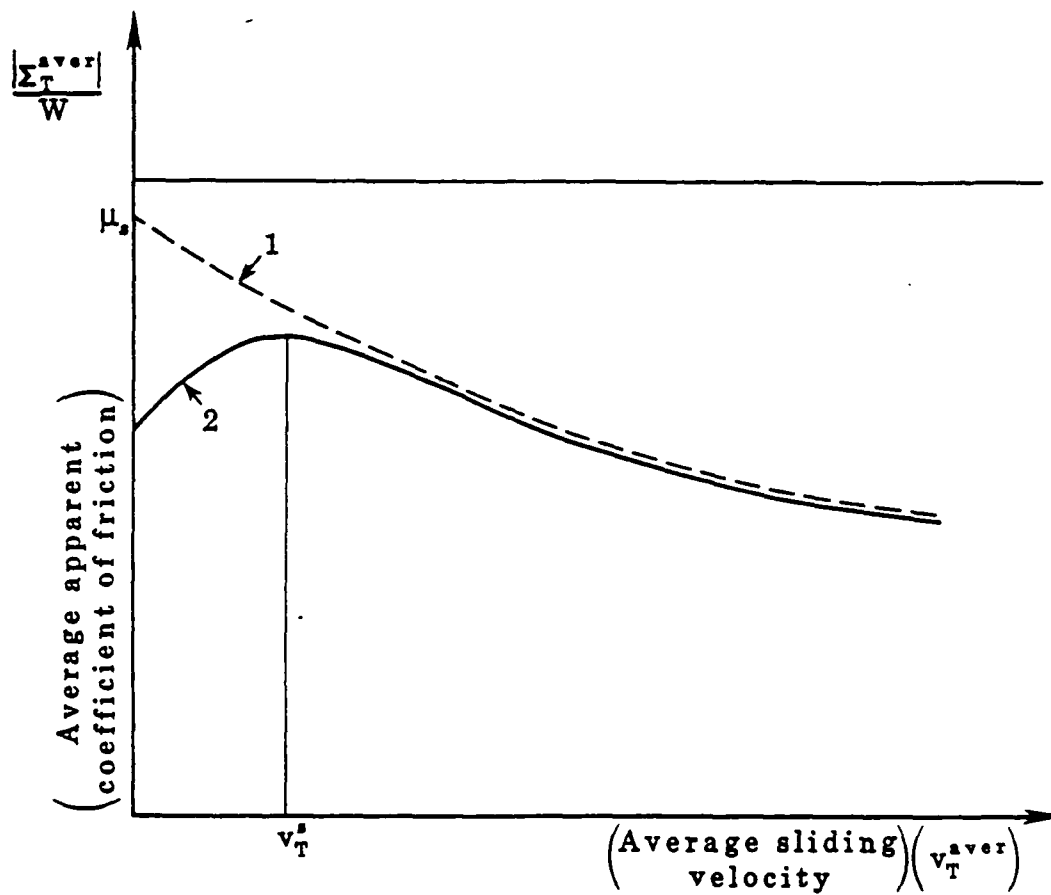


Figure 2.5: Average apparent coefficient of friction as a function of average sliding velocity according to Budanov, Kudinov and Tolstoi [18]: (1) decrease of kinetic friction owing to asymmetry of normal contact oscillations (c.f. Eq. (2.1) and Fig. 2.4); (2) increase of kinetic friction at small average sliding speeds owing to high-frequency stick-slip motion. μ_s denotes the coefficient of static friction and v_T^s denotes the average sliding velocity above which the normal oscillations are insufficient to produce any stick state.

We emphasize that the difficulty does not consist of obtaining a model of normal interface response consistent with experimental observations, and that may originate asymmetric normal oscillations. In our previous paper (Oden and Martins [56]) we used one such model which we will use again in later sections of this report. The difficult issue is: *what phenomenological friction law can simulate the effect that, according to Tolstoi and coworkers, the shocks between opposing asperities have in increasing the amplitude of the normal oscillations with the increase of sliding velocity while, simultaneously, fulfilling the average requirement of 2.1?* This is, in our view, the major open question that emerges from the work of Tolstoi and coworkers.

3 A Model of Interface Response

It results from the above exposition that, despite the extensive experimental literature devoted to the subject and the numerous theoretical interpretations proposed, a clear understanding of frictional sliding phenomena at small speeds has not been yet achieved and, of course, no model proposed has gained general acceptance.

In a previous paper [56], we developed a simple model of interface response which takes into account the normal interface deformation and which assumes that no distinction between coefficients of static and kinetic friction exists. With such a simple model we were able to simulate apparently smooth sliding motions with apparent coefficients of kinetic friction smaller than the static one and also stick-slip oscillations. The nonsymmetry of frictional contributions to the equations of motion was shown to be responsible for the dynamic instability of steady sliding equilibrium positions and this produces normal oscillations which in turn may lead to the above mentioned tangential behaviors. Such numerical results, of course, gave further support and physical insight into essential aspects of the ideas of Tolstoi on the importance of normal degrees of freedom in sliding friction phenomena.

In the present study we pursue the same line of inquiry to investigate the role played by the various physical and geometric governing parameters on the qualitative behavior of simple sliding systems. To start with, we briefly describe the model of interface response adopted and briefly discuss its major assumptions and limitations. For further details see Oden and Martins [56]. For simplicity and since it is sufficient for our purposes in this report we restrict ourselves to two-dimensional model problems.

Let us denote by σ_n and σ_T the (nominal) *normal* and *tangential stresses* at each macroscopic point on the contact interface between two metallic bodies. Let us denote by a the

penetrating approach (cf. Oden and Martins [56] and Back, et. al. [4]) and by v_T the relative sliding velocity of the interface, respectively.

The normal interface response is characterized by the constitutive equation

$$-\sigma_n = c_n a^{m_n} + b_n a^{l_n} \dot{a} \quad (3.1)$$

where $(\dot{})$ denotes differentiation with respect to time; the tangential behavior of the interface is characterized by the frictional law

$$-\sigma_T \in c_T a^{m_T} \operatorname{sgn}(v_T) \quad (3.2)$$

where $\operatorname{sgn}(\cdot)$ denotes the multivalued relation

$$\xi \in \mathbf{R} \rightarrow \operatorname{sgn}(\xi) = \begin{cases} 1 & \text{if } \xi > 0 \\ [-1, 1] & \text{if } \xi = 0 \\ -1 & \text{if } \xi < 0 \end{cases} \quad (3.3)$$

In (3.1) and (3.2), $c_n, m_n, b_n, l_n, c_T, m_T$ are interface material parameters to be determined experimentally. (We remark that in parallel work we have developed micromechanical models of metallic interfaces that can allow one to compute quantitative estimates of these parameters in terms of mechanical and geometrical properties of contact interfaces.)

The following remarks provide an explanation and interpretation of these relations:

1. The interface constitutive equation (3.1) combines a nonlinear power-law elastic contribution, $c_n a^{m_n}$, with a nonlinear dissipative component given by $b_n a^{l_n} \dot{a}$.
2. The form of the nonlinearly elastic contribution is consistent with experimental observations on the behavior of metallic interfaces subjected to low nominal pressures ($|\sigma_n| < 5\text{MPa}$) characteristic of sliding interfaces. Tables with experimental values of the constants c_n and m_n for several combinations of materials and surface finishes can be found in Back, et. al. [4].
3. The nonlinear dissipative term $b_n a^{l_n} \dot{a}$ is designed to model, only in an approximate manner, the hysteresis loops that result from the actual elasto-plastic behavior of the interface asperities. Indeed, the constitutive equation (3.1) allows for the approximation of loading paths of the form presented in Part (a) of Fig. 3.1 by loops of the form in Part (b) of Fig. 3.1. Thornley, et al [73] obtained experimentally loops of the type depicted in Part (a) of Fig. 3.1 when the surfaces were allowed to unload completely.

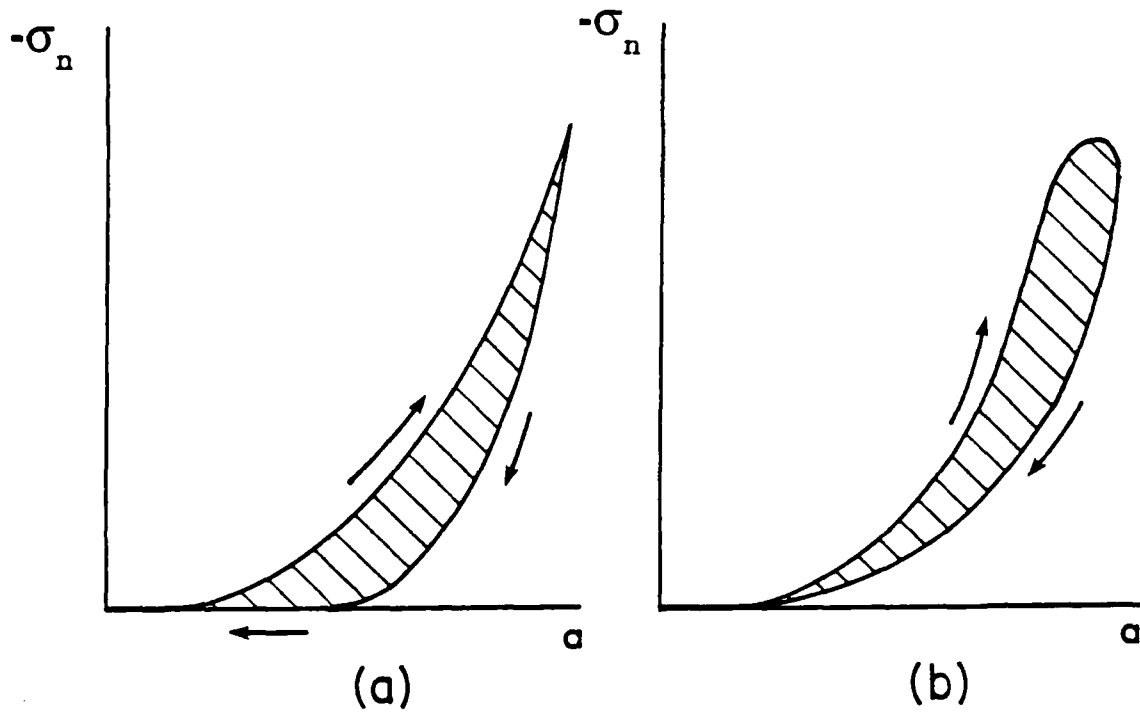


Figure 3.1: Hysteresis loops for the normal deformation of the interface (schematic). (a) Experimentally observed loop, under-static loading conditions. (b) Hysteresis loop modeled by the constitutive equation (3.1) under dynamic loading conditions.

that is to say, when some, even small, tangential reorientation of the surfaces was allowed. The idea of a similar approximation was proposed by Hunt and Crossley [40] for vibroimpact phenomena involving macroscopic Hertzian contacts. For small energy losses, the correlation between the damping coefficient b_n and the energy loss per cycle of contact is readily obtainable ([40]).

4. The friction law (3.2) is a slightly generalized local form of classical dry friction laws. It allows for possible deviations from the Amontons laws, i.e., a possible dependence of the coefficient of friction (μ) on the normal stress according to

$$\mu = C|\sigma_n|^\alpha \text{ with } \alpha = m_T/m_n - 1 \text{ and } C = c_T/c_n^{m_T/m_n}$$

if the normal dissipative effects are negligible ($b_n \approx 0$). If $m_T = m_n$ (and again, $b_n \approx 0$) the usual Coulomb's law of friction is recovered with $\mu = c_T/c_n$.

5. In using the friction law (3.2) we assume that there exists no distinction between coefficients of static and kinetic friction and no variation of the latter with the sliding speed. We introduce these assumptions because:
 - (a) The experimental observations of Tolstoi and coworkers summarized in Section 2.6 clearly suggest that at small sliding speeds those distinctions or variations are not intrinsic properties of the interfaces.
 - (b) Occurrence of friction-induced oscillations even for sliding speeds in regions where the slope of the $\mu - v_T$ curve is positive have already been observed [87]. Furthermore, instability of steady-sliding and occurrence of self-excited oscillations with some systems that have two or more degrees-of-freedom and particular geometric configurations have been explained without recourse to the classical assumption of a decreasing $\mu - v_T$ curve [3,30,31,43,67,72].

We wish thus to study how much of the frictional behavior observed at low speeds can be explained and numerically simulated without invoking those classical assumptions. The above assumptions also make it clear that neither creep nor thermal softening effects will be taken into account in the present study.

6. Dissipative effects on a metallic interface are associated mainly with plastic deformation of the interface asperities. Of course, such plastic deformation involves both

tangential and normal motions coupled in a complex manner (cf. Bowden and Tabor [14]) which we cannot expect to reproduce in detail with the dissipative terms in (3.1) and (3.2). Related to this is also the fact that the friction law (3.2) does not take into account preliminary (plastic) tangential micro-displacements known to occur before gross-sliding (Courtney-Pratt and Eisner [26]). We believe, however, that these approximations are acceptable because we are not interested in studying the details of small quasi-static evolutions of the bodies involved, but rather gross motions and oscillations for which the major contributions of the interface are its normal nonlinear elasticity and its tangential frictional dissipation, with a comparatively much smaller contribution of the normal dissipation. In this context we remark that in all the calculations done by Tolstoi and coworkers to analyze their experimental results, only the two major contributions mentioned above were taken into account, although, of course, those authors were perfectly aware of the existence of the normal dissipation.

7. The constitutive assumptions (3.1) and (3.2) are not completely consistent with some aspects of the work of Tolstoi and coworkers:

- (a) The nonlinearly elastic contribution in (3.1) is not of the form experimentally observed by those authors for the surfaces and loads they worked with, but, within the range of validity of our model, it leads to normal microvibrations with frequencies close to those considered by Tolstoi as typical: 10^3Hz . Indeed, denoting by W the weight of a body, by A the nominal area of its nominally flat contact support surface, it results from (3.1) that in the equilibrium position a_o (recall Fig. 2.4)

$$W = c_n A a_o^{m_n}.$$

The (linearized) normal stiffness of the support at that equilibrium position is

$$K_{on} = m_n c_n A a_o^{m_n-1} = m_n W / a_o$$

so that, with g denoting acceleration due to gravity, the frequency of the small free normal oscillations is

$$f_{on} \approx \frac{1}{2\pi} \sqrt{\frac{K_{on}}{W/g}} = \frac{1}{2\pi} \sqrt{\frac{m_n g}{a_o}}.$$

With typical values of m_n in the range of 2 to 3, with $g \approx 10ms^{-2}$ and with values of a_o in the range of 0.3 to $10\mu m$ (see Figs.1(a), (b), (c) in Back, Burdekin and Cowley [4]), the frequency of the free normal oscillation will be somewhere in the range of $2 \times 10^2 Hz$ to $2 \times 10^3 Hz$.

- (b) The phenomenological laws (3.1) and (3.2) are not consistent with the asymmetry mechanism (I) in Section 2.6. Indeed, taking, for simplicity, $m_T = m_n$, $\mu = \mu_s = c_T/c_n$ in (3.2), and assuming the occurrence of a sliding motion accompanied by steady normal oscillations and with an average sliding velocity sufficiently large that no stick state occurs ($|v_T^{aver}| > v_T^s$ in Fig. 2.5), it results from (3.2) that the average friction force is equal to the product of the coefficient of (static) friction with the average normal contact force, i.e.,

$$|\Sigma_T^{aver}| = \mu_s |\Sigma_n^{aver}| = \mu_s W,$$

which is clearly different from (2.1). We do not address here what we called in the previous section the major open question resulting from Tolstoi's work. However, we will be able to model their high-frequency stick-slip mechanism (II) and also other effects that are attributed by those authors to their mechanism (I).

8. Finally, we observe that no time or rate dependence of the static friction are considered with the law (3.2). Consequences of this on our results will be analyzed and, much in the spirit of Tolstoi's ideas on the apparent reductions of static friction, a preliminary study on the effect of external perturbations on the measurable static friction will be conducted in Section 5.4.

4 Normal Oscillations And a Stick-Slip Motion With A Two-Degree-of-Freedom System

4.1 Governing Equations

In this section we study the effect of normal oscillations on frictional sliding by using a simple two degree-of-freedom model of a body of mass M and weight W , restrained by a horizontal spring (K_x) and sliding with friction on a surface which moves with a prescribed time independent tangential velocity \dot{U}_x^C (see Fig. 4.1). Assuming that the interface laws hold, with $b_n \equiv 0$, on the contact surface Γ_C (of nominal area A), the motion $t \in [0, T] \rightarrow \mathbf{u}(t) = [u_x(t), u_y(t)]^t \in \mathbf{R}^2$, $T > 0$, satisfies, for *a.e.* t ,

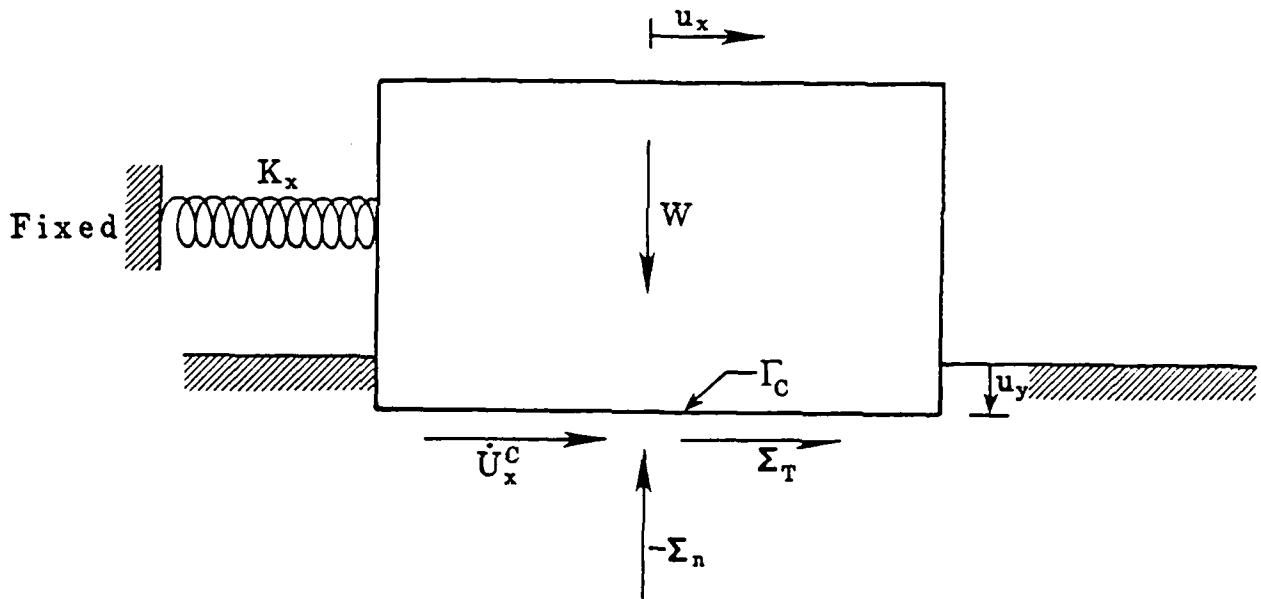


Figure 4.1: Two degree-of-freedom model of a block sliding with friction on a moving foundation; u_x = tangential displacement; u_y = normal (penetrating) displacement; \dot{U}_x^C = driving velocity; K_x = tangential stiffness; W = weight of the body; Σ_n = normal contact force; Σ_T = friction force.

$$M\ddot{u}_x(t) + K_x u_x(t) \in -c_T A[(u_y(t))_+]^{m_T} \text{sgn}(\dot{u}_x(t) - \dot{U}_x^C) \quad (4.1)$$

$$M\ddot{u}_y(t) + c_n A[(u_y(t))_+]^{m_n} = W \quad (4.2)$$

with the initial conditions

$$u_x(0) = \bar{u}_{x0}, u_y(0) = \bar{u}_{y0} \quad (4.3)$$

$$\dot{u}_x(0) = \bar{u}_{x1}, \dot{u}_y(0) = \bar{u}_{y1}. \quad (4.4)$$

A nondimensional form of these equations and inclusion is

$$u_x''(\tau) + u_x(\tau) \in -f \cdot [(u_y(\tau))_+]^{m_T} \cdot \text{sgn}(u_x'(\tau) - U_x^C) \quad (4.5)$$

$$s u_y''(\tau) + [(u_y(\tau))_+]^{m_n} = 1 \quad (4.6)$$

$$u_x(0) = \bar{u}_{x0}, u_y(0) = \bar{u}_{y0} \quad (4.7)$$

$$u_x'(0) = \bar{u}_{x1}, u_y'(0) = \bar{u}_{y1}. \quad (4.8)$$

Here, τ denotes the nondimensional time, $\tau = \omega t$, $\omega = (K_x/M)^{1/2}$; $(\cdot)'$ denotes differentiation with respect to τ ; $u_x(\tau) = (K_x/W) \cdot u_x(t)$ and $u_y(\tau) = (c_n A/W)^{1/m_n} u_y(t)$ denote the nondimensional displacements; f denotes the *friction parameter*.

$$f = \frac{c_T}{c_n^{m_T/m_n}} \left(\frac{W}{A} \right)^{\frac{m_T}{m_n} - 1};$$

s denotes a *stiffness parameter* which measures the stiffness of the tangential spring relatively to the (linearized) normal stiffness K_{on} ,

$$s = \frac{K_x}{W} \left(\frac{W}{c_n A} \right)^{1/m_n};$$

$U_x^C = \dot{U}_x^C \cdot (MK_x)^{1/2}/W$ is the nondimensional driving velocity; and $\bar{u}_{x0}, \bar{u}_{y0}, \bar{u}_{x1}, \bar{u}_{y1}$ denote now the initial nondimensional displacements and velocities.

For computational purposes, the multivalued relation $\text{sgn}(\cdot)$ is approximated by the function $\phi_\varepsilon(\cdot)$ defined by

$$\xi \in \mathbf{R} \rightarrow \phi_\varepsilon(\xi) = \begin{cases} 1 & \text{if } \xi > \varepsilon \\ (2 - \frac{|\xi|}{\varepsilon}) \frac{\xi}{\varepsilon} & \text{if } |\xi| \leq \varepsilon \\ -1 & \text{if } \xi < -\varepsilon \end{cases} \quad (4.9)$$

where the regularization parameter ε is a "small" positive real number. The differential inclusion (4.5) is then approximated by the equation

$$u_x''(\tau) + u_x(\tau) = -f[(u_y(\tau))_+]^{m_T} \cdot \phi_\varepsilon(u_x'(\tau) - U_x'^C) \quad (4.10)$$

and the system of equations (4.6) and (4.10) with initial conditions (4.7) and (4.8) is solved numerically using Newmark's method as described in Oden and Martins [56].

4.2 Numerical Results

The results reported here were obtained with $f = 0.3$ and $m_n = m_T = 2$. Since $m_n = m_T$ and $b_n \equiv 0$, the friction law (3.2) coincides with the usual Coulomb's law and the friction parameter f coincides with the usual (independent of the normal pressure) coefficient of friction $\mu (f = c_T/c_n = \mu)$. The nondimensional initial conditions for the motion are the following: $\bar{u}_{x0} = 0$; $\bar{u}_{x1} = U_x'^C$; $\bar{u}_{y0} = 0.5$; $\bar{u}_{y1} = 0$, i.e., the body starts its motion stuck with the driving surface and away from the nondimensional normal equilibrium position $u_y = 1$. As a result of the latter, normal free oscillations of the body will occur, governed by the nonlinear dynamic equation (4.6). The normal oscillation of the body produces variations on the maximum available nondimensional friction force $f[(u_y)_+]^2$ which, in turn, are responsible for stick-slip tangential oscillations. Note that, if no normal oscillations existed, the body, after the initial stick phase, would continue to slide, undergoing a free sinusoidal oscillation. In the first example (Fig. 4.2), the stiffness parameter is $s = 3$ and the driving velocity $U_x'^C = 0.3286$. In Fig. 4.2 we present a plot of the resulting stick-slip motion; in Fig. 4.3 the time evolution of the nondimensional normal force $-\Sigma_n/W = [(u_y)_+]^2$ and the nondimensional regularized friction force

$$\Sigma_T/W = \mu \cdot [(u_y)_+]^2 \cdot \phi_\varepsilon(u_x' - U_x'^C)$$

are shown; finally in Fig. 4.4 the evolution of the ratios Σ_T/Σ_n and Σ_T/W is plotted against the corresponding nondimensional sliding velocities $(u_x' - U_x'^C)$.

In the second (Figs. 4.5 to 4.7) and third (Figs. 4.8 and 4.9) examples, the stiffness parameter is, in both cases, $s = 1.911$ and the driving velocity is $U_x'^C = 2.623 \times 10^{-1}$ and $U_x'^C = 8.743 \times 10^{-2}$, respectively.

4.3 Discussion

The numerical results presented here show clearly that oscillations in the direction normal to the sliding plane may induce a stick-slip sliding. However, this is not at all surprising

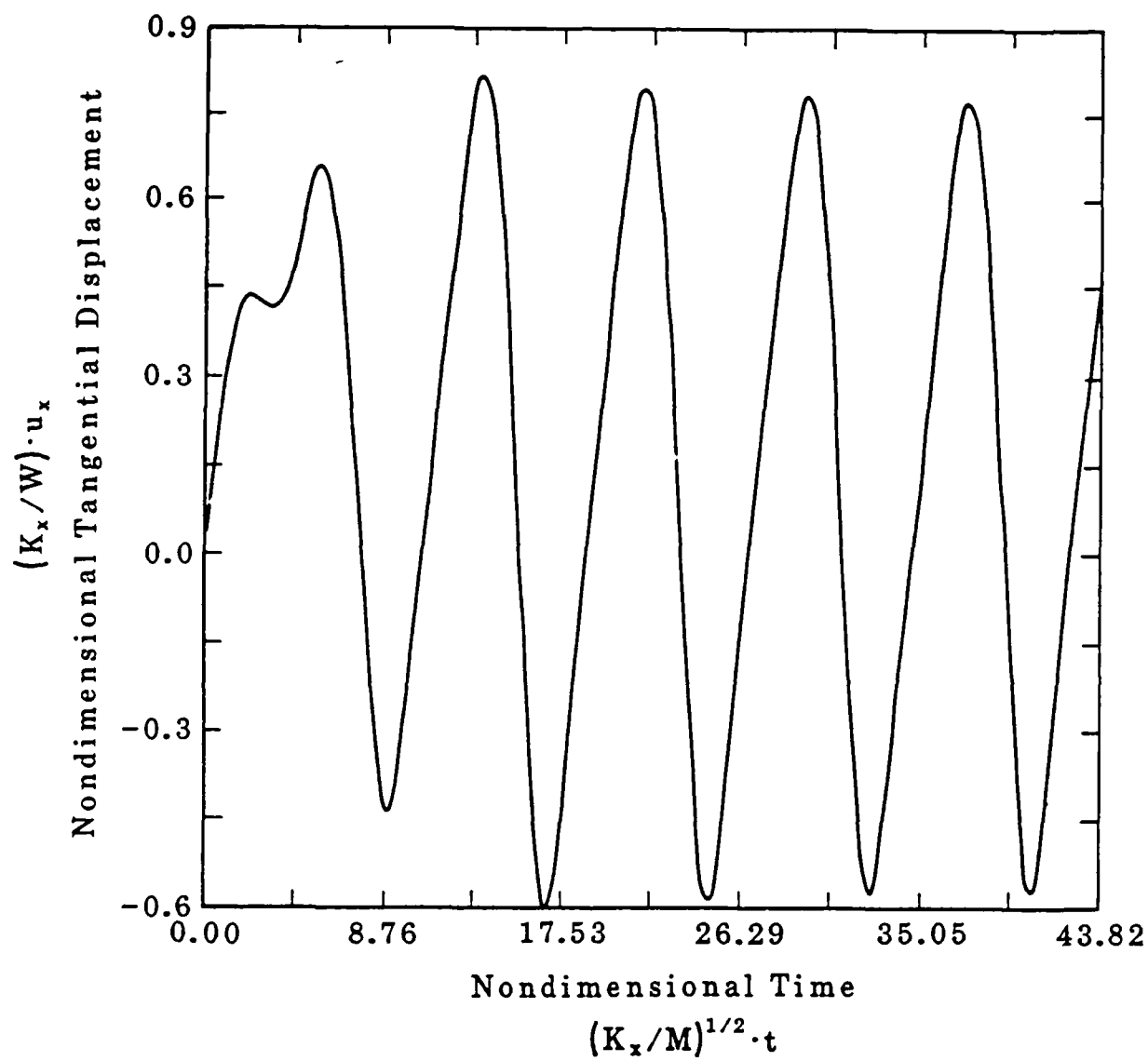


Figure 4.2: Stick-slip motion induced by normal oscillations ($s = 3$, $U_r^C = 0.3286$). Tangential displacement vs. time.

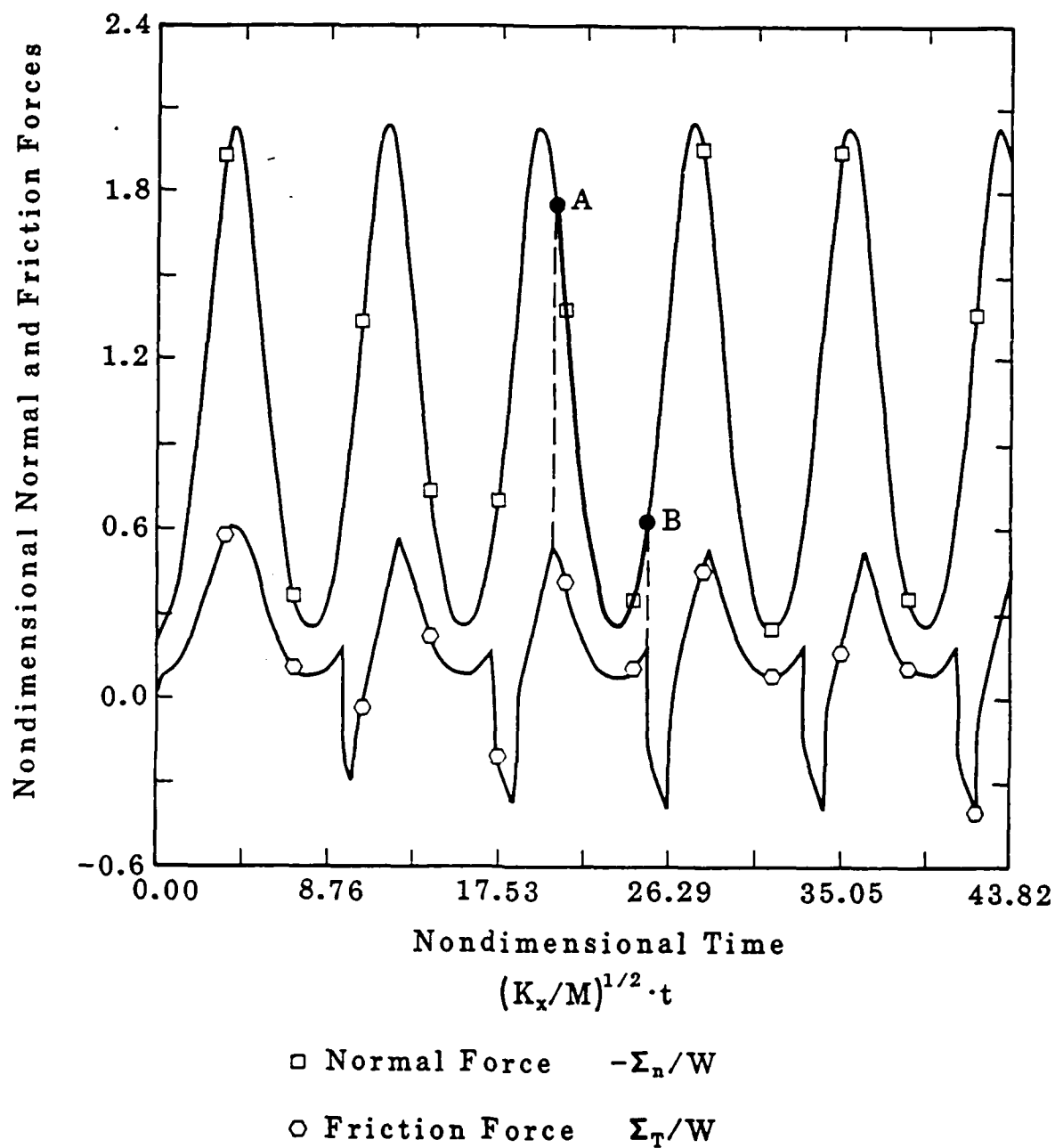


Figure 4.3: Stick-slip motion induced by normal oscillations ($s = 3$, $U_x^C = 0.3286$). Normal and friction force vs. time.

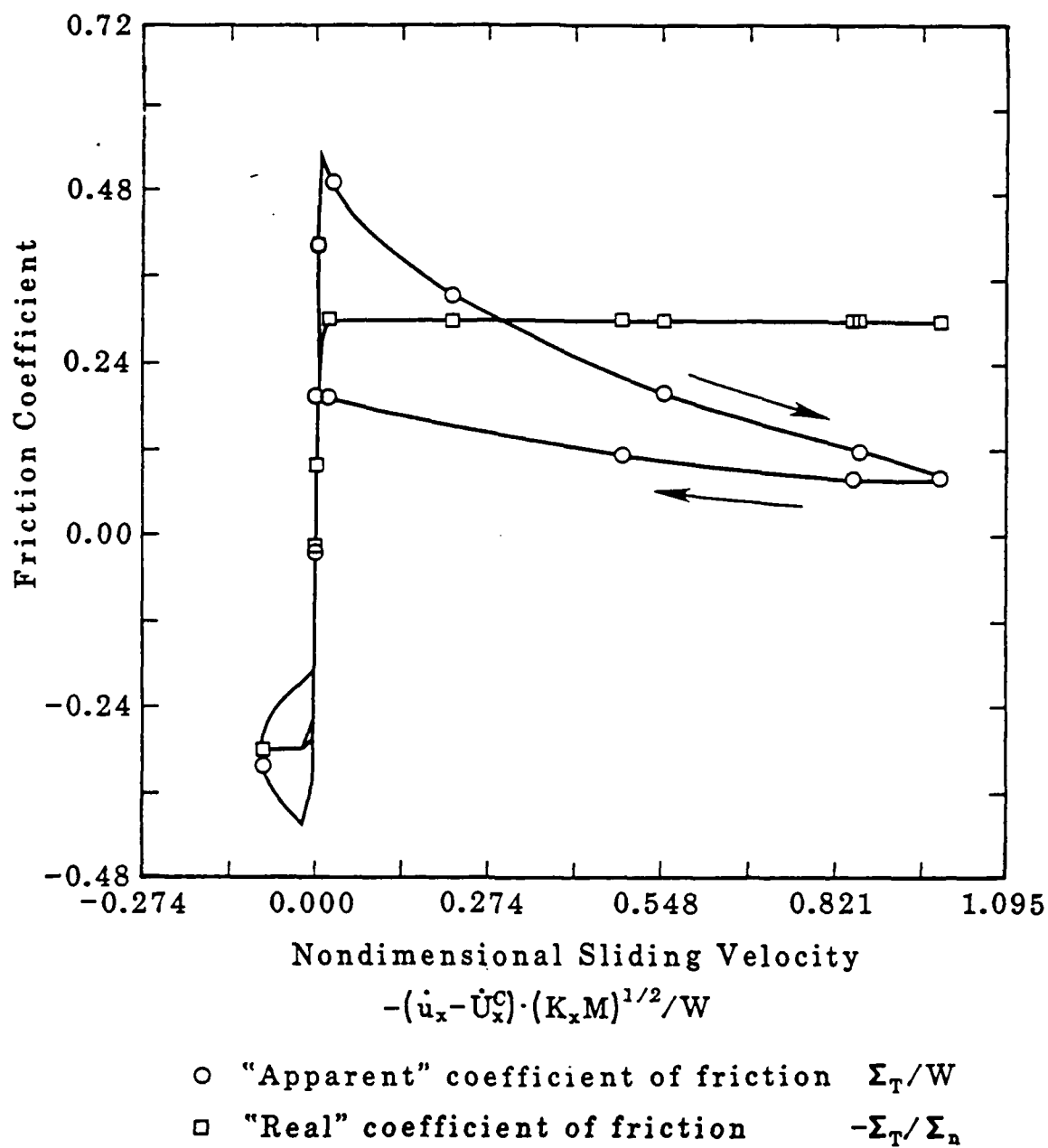


Figure 4.4: Stick-slip motion induced by normal oscillations ($s = 3$, $U_x^c = 0.3286$). Real and apparent coefficient of friction vs. sliding velocity.

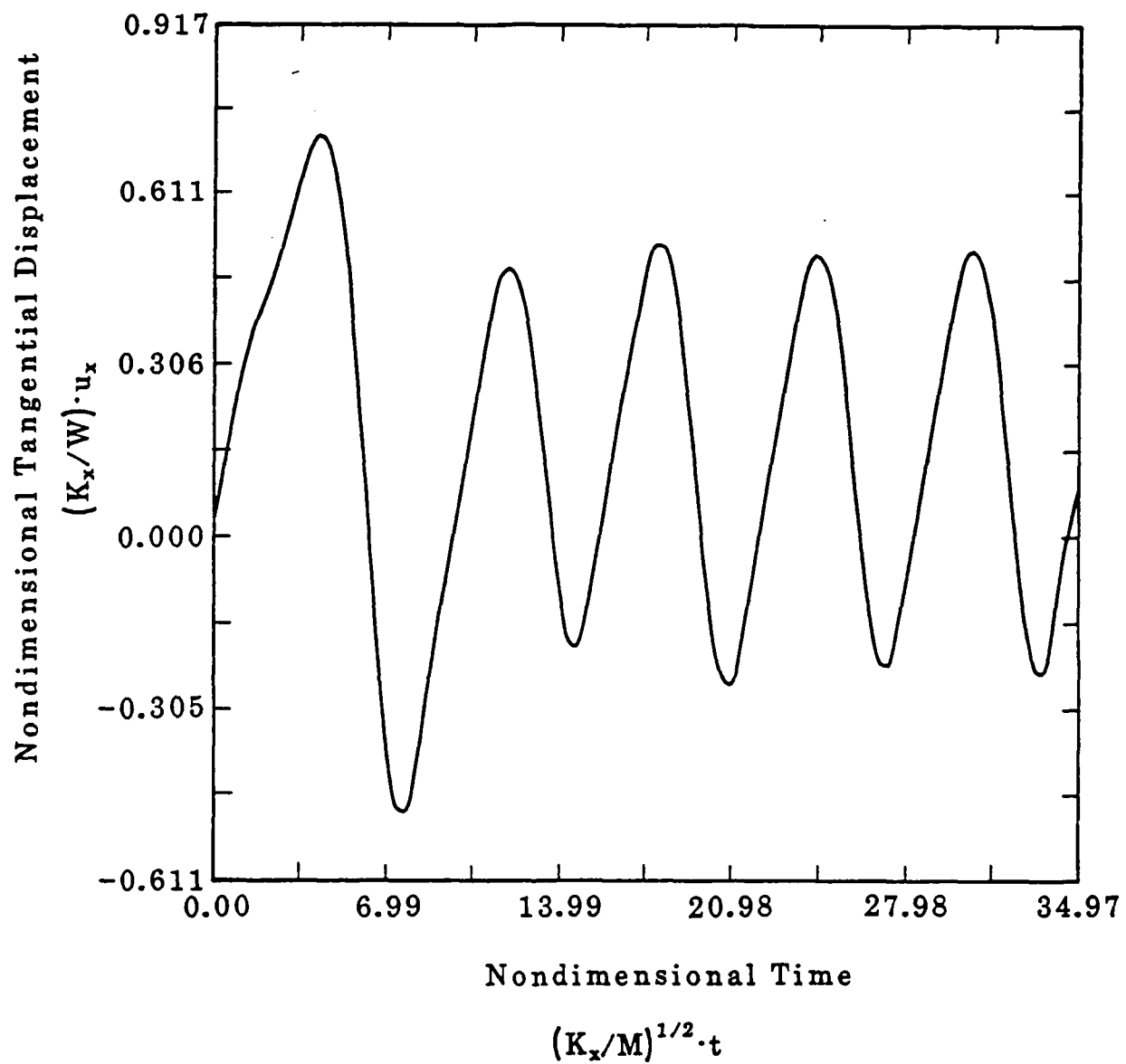


Figure 4.5: Stick-slip motion induced by normal oscillations ($s = 1.911$, $U_r^C = 0.2623$). Tangential displacement vs. time.

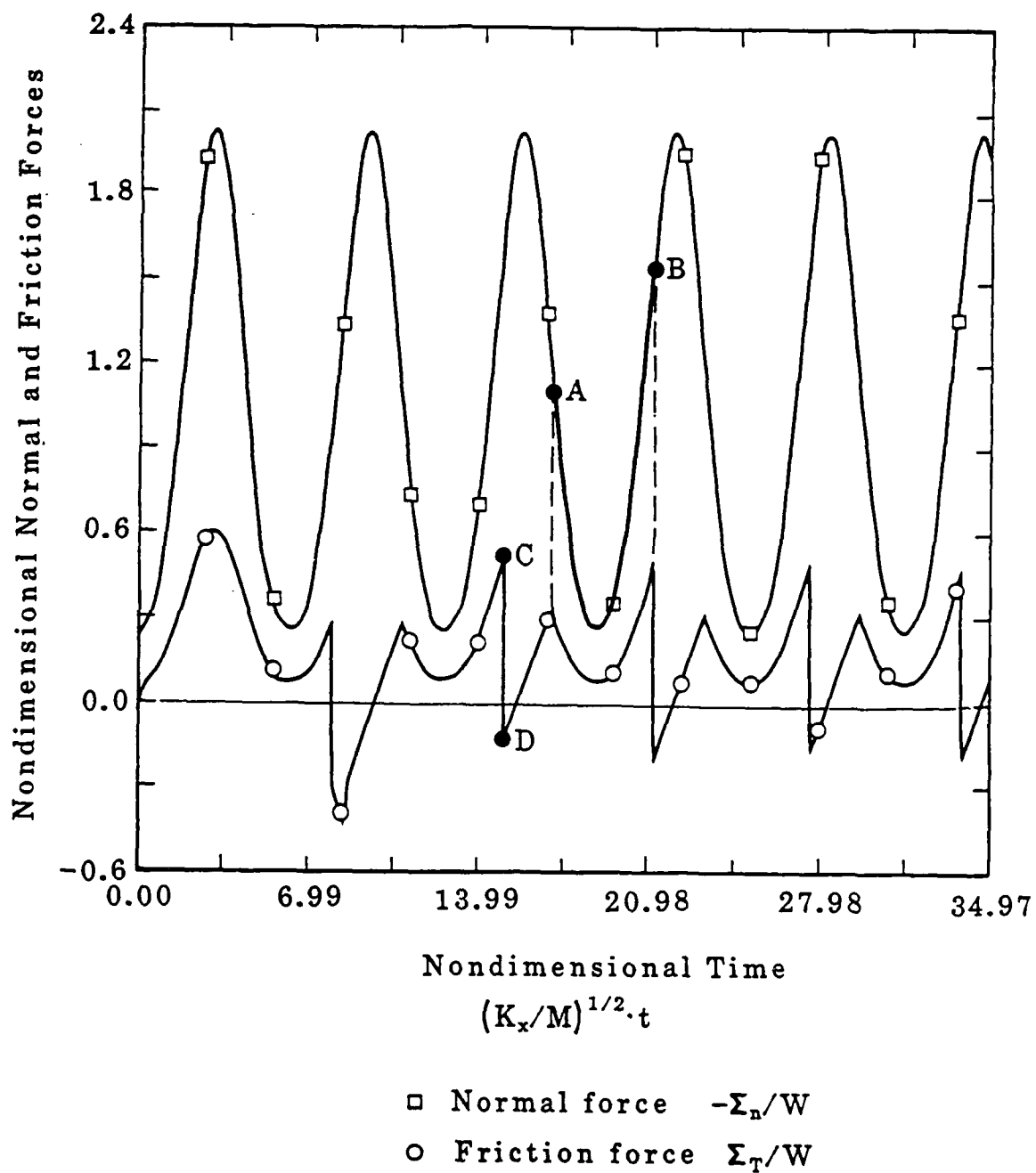


Figure 4.6: Stick-slip motion induced by normal oscillations ($s = 1.911$, $U_x^C = 0.2623$). Normal and friction force vs. time.

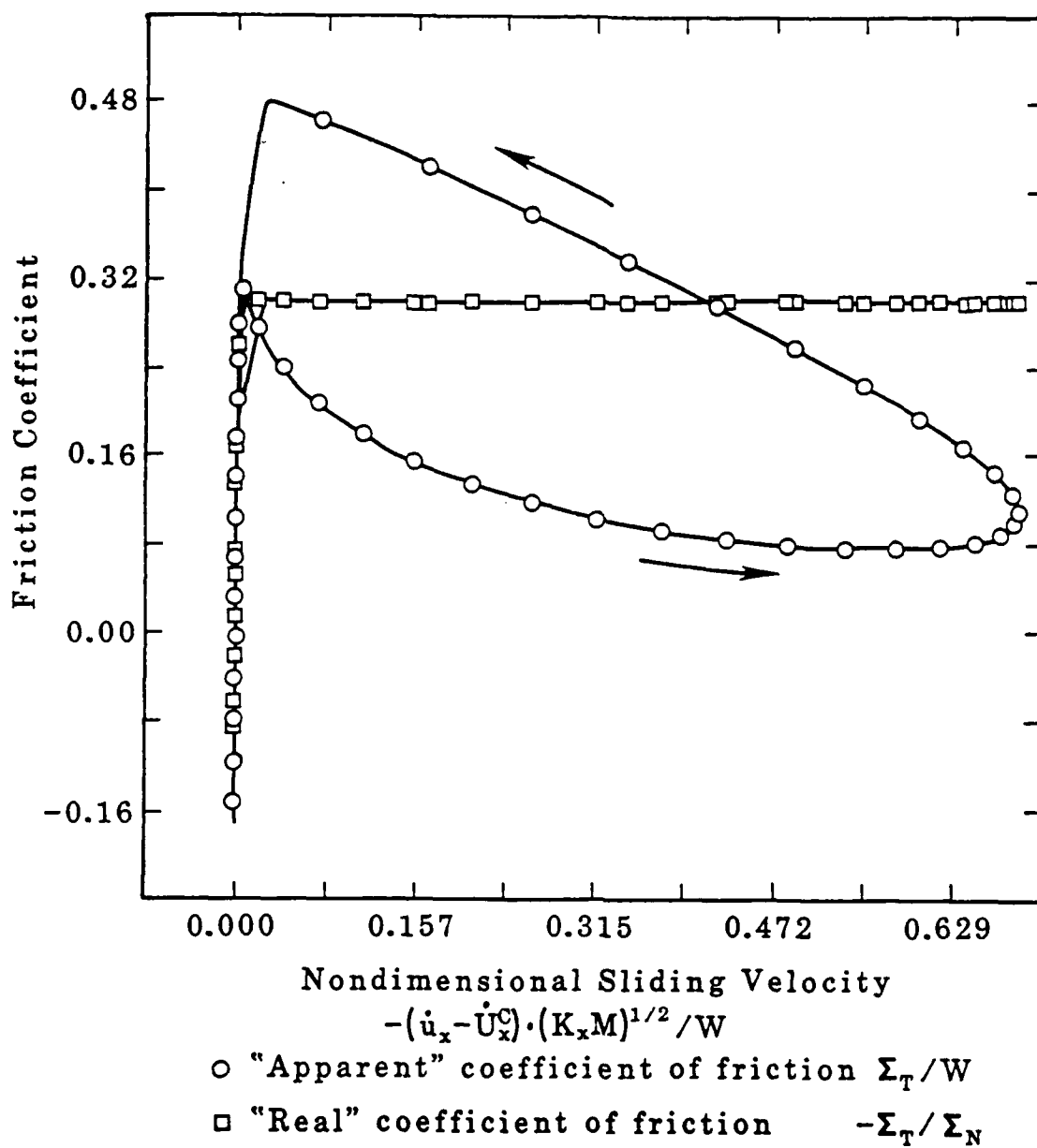


Figure 4.7: Stick-slip motion induced by normal oscillations ($s = 1.911$, $U_x^{IC} = 0.2623$). Real and apparent coefficient of friction vs. sliding velocity.

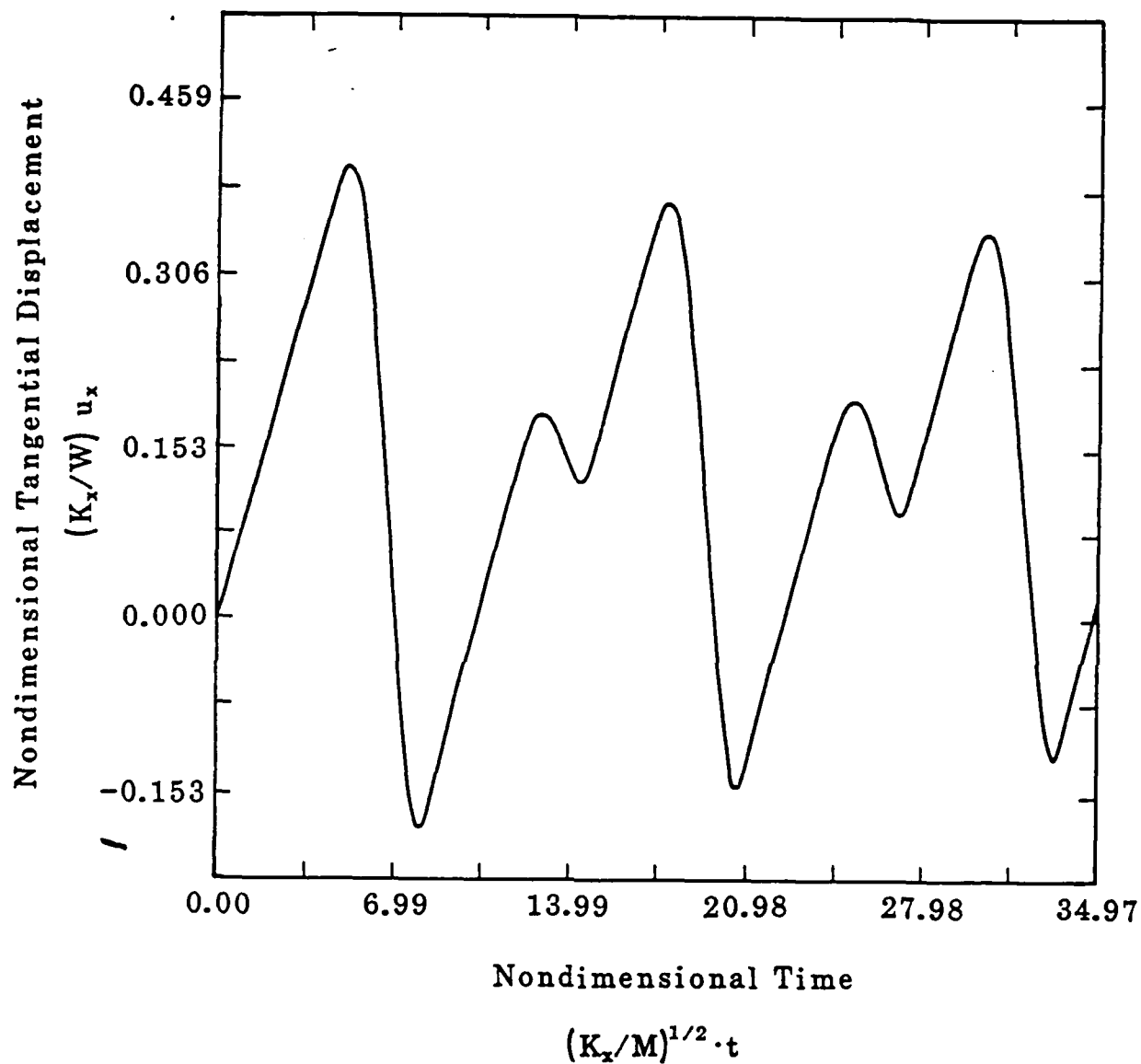


Figure 4.8: Stick-slip motion induced by normal oscillations ($s = 1.911$, $U_x^C = 8.743 \times 10^{-2}$). Tangential displacement vs. time.

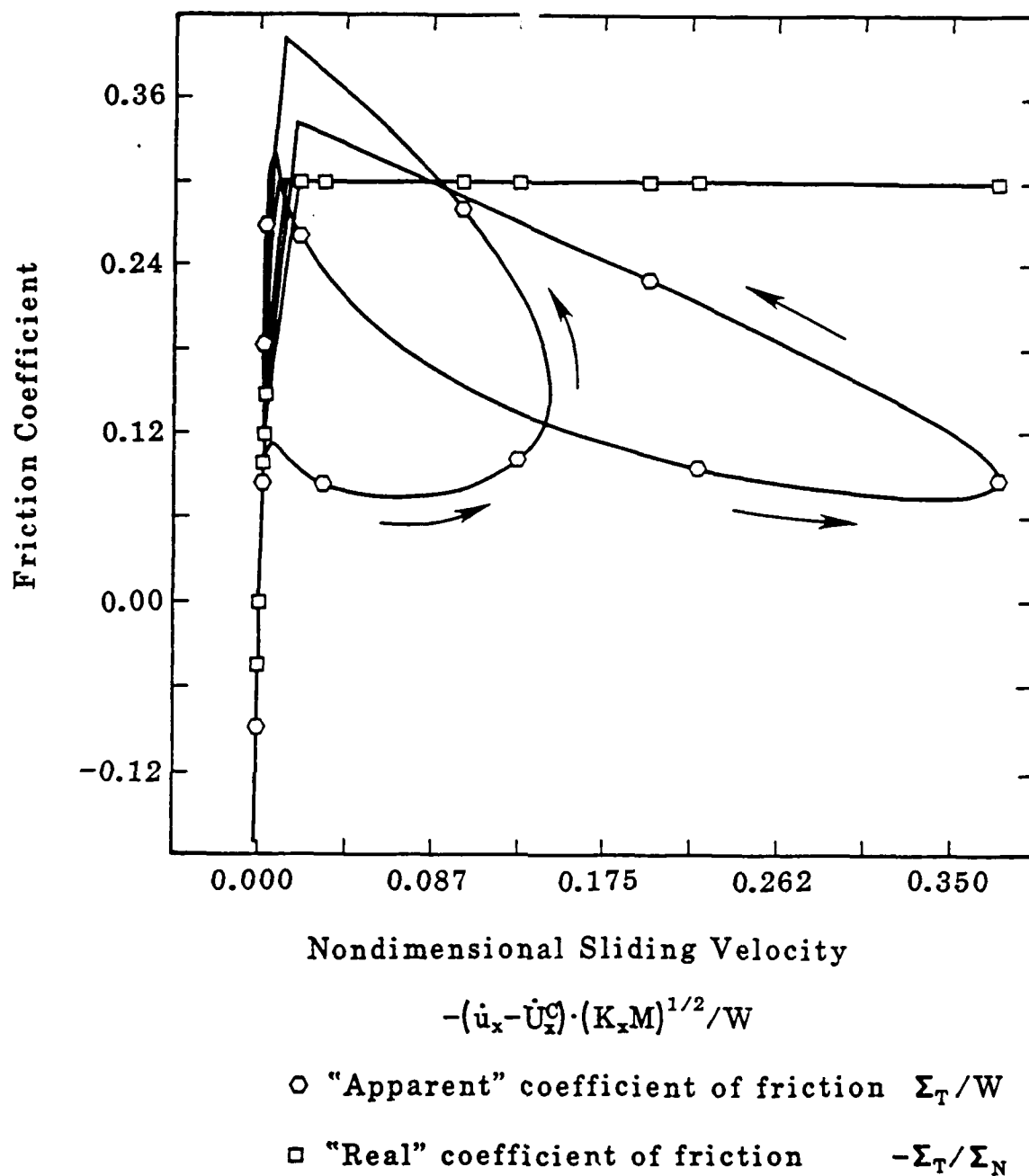


Figure 4.9: Stick-slip motion induced by normal oscillations ($s = 1.911$, $U_x^{IC} = 8.743 \times 10^{-2}$). Real and apparent coefficient of friction vs. sliding velocity.

and cannot explain actual stick-slip relaxation oscillations *unless some justification is found for the occurrence and persistence of the normal oscillation*. With the simple model used here, any amount of normal damping would damp out the normal oscillation and the stick-slip oscillations would cease. What we do here is not completely new, either: Boyer and Cameron [85] produced earlier similar stick-slip traces induced by normal motions. Those authors used a statistical model for the normal deformability of the interface, while we use the phenomenological normal interface law (3.1) with the same purpose; those authors assumed a linear relationship between the decrease of normal penetration and the increase of sliding velocity, while here we simply set the slider in a normal free oscillation.

What we believe to be interesting in our results is the fact that, although we always assume a *real* coefficient of friction independent of the sliding velocity, the *apparent* coefficient of friction (instantaneous friction force/nominal normal load = Σ_T/W) describes loops qualitatively very similar to those experimentally found by various authors. The clockwise loop of Fig. 4.4 resembles those obtained by, among others, Bell and Burdekin [5], and the counterclockwise loops of Figs. 4.7 and 4.9 resemble those obtained by Antoniou, Cameron and Gentle [1]. We observe that the different types of loops in Figs. 4.4, 4.7, and 4.9, result from varying the (arbitrarily assumed) values of s and U_x^C and this is equivalent to assuming *different variations of normal force during the sliding phase of the stick-slip motion*: Compare the portions AB of the plots of the normal force in Figs. 4.3 and 4.6. For actual stick-slip motions the form (and orientation) of the loops described by the apparent coefficient of friction will be thus very sensitive to the forms of the associated normal jumps, and these certainly will depend not only on the interface materials considered, but also on the dynamic properties of the experimental apparatus used.

One should also observe that the shape and orientation of the coefficient of friction-sliding speed loops that are most commonly reported in the literature [5, 47, 81] are those of the type in Fig. 4.4. As seen in Fig. 4.3, this corresponds to a normal force that sharply decreases upon the initiation of sliding and which remains small during the whole slip phase. Clearly, this is consistent with Tolstoi's observations and interpretations on the origin of stick-slip motions.

From these observations we conclude that *plots of the variation of the coefficient of friction with the sliding velocity, obtained experimentally in the course of stick-slip oscillations, are essentially meaningless, if corresponding variations of the normal contact force are not appropriately taken into account and if these variations are significant, as suggested by the experimental results summarized in Section 2.6.*

Remark 4.3.1. Actually, Antoniou, Cameron and Gentle [1] claim that their coefficient of friction-sliding velocity plots were obtained by taking into account instantaneous values of the normal contact force and not by merely dividing the instantaneous friction force by the nominal normal load. It is not clear, however, how reliable their computation of the normal contact force is, since they assumed it solely dependent on the normal velocity (which does not appear to have any possible physical justification) and since no data or study of the normal oscillation is presented. Furthermore, the tangential and normal displacements of the slider are not measured directly, but are inferred from the readings of two strain gauges placed in two *different* positions on the flexible arm that supports the slider. The accuracy of the predicted displacements of the slider (especially the very small normal displacement) seems thus questionable. In particular, by comparing our computed loops with those calculated from the experimental data, we suspect that the jump in the friction coefficient predicted by Antoniou *et al.* to occur at a sliding velocity different from zero (at point C of their Fig. 8, p.246, op. cit.) is nothing but the result of misplacing the end of the slip phase at which a discontinuity of the friction force and the tangential acceleration occurs (see, for example, the discontinuity CD of the friction force in our Fig. 4.6). We note that such a jump has been observed experimentally (and correctly located) by other authors (e.g. Fig. 9, p. 555 [5] and Fig. 3, p.181 [47]) who measured experimentally the tangential acceleration of the slider.

Of course, we do not claim that the simple two-degree-of-freedom model employed in this section can be used, without further coupling between normal and tangential freedoms, to model actual stick-slip oscillations. The normal motion governed by equation (4.2) is not affected at all by the tangential motion and this is not consistent with the experimental observations summarized in Section 2.6.

A simple mechanical system that can capture the coupling between normal and tangential freedoms, is a rigid body in plane motion. We shall employ it in the studies of the remaining sections of this report.

Remark 4.3.2. For two degree-of-freedom systems, additional coupling between normal and tangential freedoms may exist if, for example, some restraining spring is inclined relatively to the plane of sliding. This situation has been considered in different forms by some authors, in particular by Shobert [67], and Spurr [72] (see also the discussion of the same paper by Rabinowicz) in the context of stability of steady-sliding and occurrence of frictional oscillations; and by Klarbring [83] in the context of nonexistence or nonuniqueness

of solutions to quasistatic evolution friction problems. Also, for a two degree-of-freedom lubricated sliding system, additional coupling between the normal and tangential freedoms was introduced by Kudinov and Lisitsyn [82] when they considered a lifting hydrodynamic component for the normal contact force, dependent on the tangential sliding velocity. For continuum elastic systems, of course, normal and tangential displacements are necessarily coupled due to the deformability (Poisson's effect) of the contacting bodies.

5 Low Speed Dynamic Friction Phenomena With a Three-Degree-of-Freedom Rigid Body Model

5.1 Governing Equations

We consider here a rigid block (see Fig. 5.1) with dimensions $L \times H \times B$, weight W , mass M , moment of inertia with respect to the axis through the mass center $I = M(L^2 + H^2)/12$, restrained by a horizontal arm with elastic stiffness $K_x (> 0)$ and damping coefficient $C_x (\geq 0)$, and sliding with friction on a surface which moves with a prescribed tangential velocity \dot{U}_x^C . The block is supposed to have plane motion, the corresponding degrees-of-freedom being: the tangential and normal (penetrating) displacements of the center of mass $G(u_x$ and u_y , respectively) and the rotation (u_Θ), as depicted in Fig. 5.1. *The rotation u_Θ is assumed to be small so that $\sin u_\Theta \approx \tan u_\Theta \approx u_\Theta$ and $\cos u_\Theta \approx 1$.* Along the flat candidate contact surface Γ_C the contact laws (3.1) and (3.2) are assumed to hold. In view of the geometry of the present problem it follows from the normal contact equation (3.1) that the vector of generalized forces Σ_n associated with the normal stresses on Γ_C satisfies, at each time t ,

$$-\Sigma_n(t) = P(u(t)) + Q(u(t), \dot{u}(t))$$

where $u(t) = [u_x(t), u_y(t), u_\Theta(t)]^t \in \mathbf{R}^3$ is the vector of generalized displacements at time t , and

$$w \in \mathbf{R}^3 \rightarrow P(w) = Bc_n \int_{-L/2}^{L/2} [(w_y - xw_\Theta)_+]^{m_n} \begin{Bmatrix} 0 \\ 1 \\ -x \end{Bmatrix} dx \quad (5.1)$$

$$(w, v) \in \mathbf{R}^3 \times \mathbf{R}^3 \rightarrow Q(w, v) = Bb_n \int_{-L/2}^{L/2} [(w_y - xw_\Theta)_+]^{1_n} (v_y - xv_\Theta) \begin{Bmatrix} 0 \\ 1 \\ -x \end{Bmatrix} dx \quad (5.2)$$

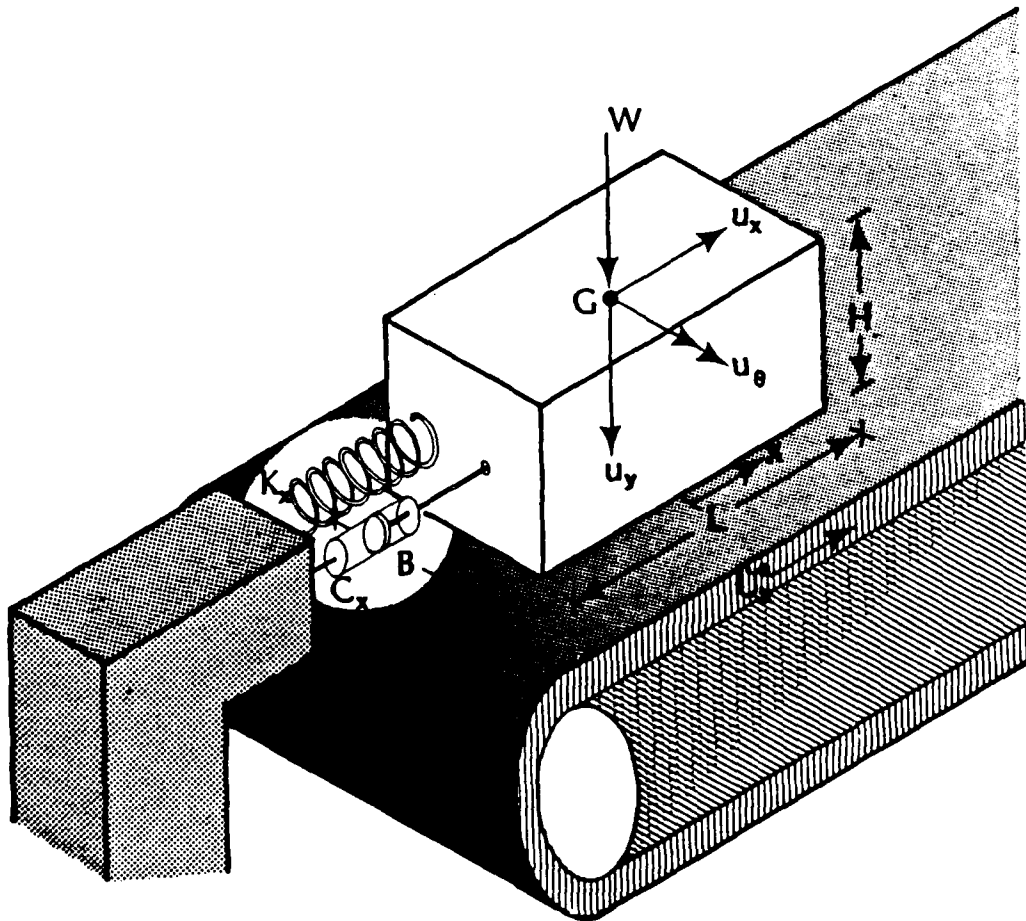


Figure 5.1: Geometry and degrees-of-freedom of a rigid block sliding with friction on a moving foundation.

On the other hand, from the friction law (3.2) it follows that the vector of generalized forces Σ_T associated with the friction stresses on Γ_C satisfies at each time t

$$-\Sigma_T(t) \in J(u(t), \dot{u}(t) - \dot{\Phi}(t))$$

where

$$(w, v) \in \mathbf{R}^3 \times \mathbf{R}^3 \rightarrow J(w, v) = B_{CT} \operatorname{sgn}(v_x + \frac{H}{2} v_\Theta) \begin{Bmatrix} 1 \\ 0 \\ \frac{H}{2} \end{Bmatrix} \int_{-L/2}^{L/2} [(w_y - x w_\Theta)_+]^{m_T} dx \quad (5.3)$$

$$t \in [0, T] \rightarrow \dot{\Phi}(t) = \begin{Bmatrix} \dot{U}_x^C(t) \\ 0 \\ 0 \end{Bmatrix}.$$

For this system the motion $t \in [0, T] \rightarrow u(t) \in \mathbf{R}^3$ has thus to satisfy, for *a.e.* t , the following differential inclusion:

$$\begin{aligned} M\ddot{u}(t) + C\dot{u}(t) + Ku(t) \\ + P(u(t)) + Q(u(t), \dot{u}(t)) \\ + J(u(t), \dot{u}(t) - \dot{\Phi}(t)) \ni F \end{aligned} \quad (5.4)$$

and, at $t = 0$, the initial conditions

$$u(0) = \bar{u}_0, \dot{u}(0) = \bar{u}_1. \quad (5.5)$$

where $M, C, K \in \mathbf{R}^9$ denote the *symmetric* mass, linear damping and linear stiffness matrices and $F(t) \in \mathbf{R}^3$ denotes the vector of applied forces at time t . The matrix M is *positive definite* and the matrices C and K are *positive semi-definite*. The form of these matrices and vector in the present problem is

$$M = \begin{bmatrix} M & 0 & 0 \\ 0 & M & 0 \\ 0 & 0 & I \end{bmatrix}; C = \begin{bmatrix} C_x & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; K = \begin{bmatrix} K_x & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; F(t) = \begin{Bmatrix} 0 \\ W \\ 0 \end{Bmatrix}.$$

In the next sections, we study conditions for the occurrence of smooth or intermittent sliding motions. For given $F \in \mathbf{R}^3$ and $\dot{U}_x^C \in \mathbf{R} - \{0\}$ both independent of time, a smooth

steady sliding of the slider in Fig. 5.1 is an equilibrium position $\mathbf{u}_0 \in \mathbf{R}^3$ satisfying (set $\dot{\mathbf{u}} = \ddot{\mathbf{u}} = \mathbf{0}$ in 5.4).

$$\mathbf{K}\mathbf{u}_0 + \mathbf{P}(\mathbf{u}_0) + \mathbf{J}_\eta(\mathbf{u}_0) = \mathbf{F} \quad (5.6)$$

where

$$\mathbf{w} \in \mathbf{R}^3 \rightarrow \mathbf{J}_\eta(\mathbf{w}) \stackrel{\text{def}}{=} -Bc_T\eta \begin{Bmatrix} 1 \\ 0 \\ H/2 \end{Bmatrix} \int_{-L/2}^{L/2} [(w_y - xw_\Theta)_+]^{m_T} dx (= \mathbf{J}(\mathbf{w}, \mathbf{O} - \dot{\Phi})) \quad (5.7)$$

and

$$\eta \stackrel{\text{def}}{=} \text{sgn}(\dot{U}_x^C) = \dot{U}_x^C / |\dot{U}_x^C|, \dot{U}_x^C \neq 0.$$

Remark 5.1.1. For every $\mathbf{v} \in N_\delta(\mathbf{0}) = \{\mathbf{v} \in \mathbf{R}^3 : |v_x| + |v_y| + |v_\Theta| < \delta = |\dot{U}_x^C| / \max(1, H/2)\}$ we have $|v_x + (H/2)v_\Theta| < |\dot{U}_x^C|$ so that $\text{sgn}(v_x + (H/2)v_\Theta - \dot{U}_x^C) = \text{sgn}(-\dot{U}_x^C) = -\eta$, and $\forall(\mathbf{w}, \mathbf{v}) \in \mathbf{R}^3 \times N_\delta(\mathbf{0}) \mathbf{J}(\mathbf{w}, \mathbf{v} - \dot{\Phi}) = \mathbf{J}_\eta(\mathbf{w})$, i.e., for (\mathbf{w}, \mathbf{v}) in a sufficiently small neighborhood in the phase space $\mathbf{R}^3 \times \mathbf{R}^3$ of the equilibrium position $(\mathbf{u}_0, \mathbf{0})$, $\mathbf{J}(\mathbf{w}, \mathbf{v} - \dot{\Phi})$ is single valued and independent of \mathbf{v} (velocity). For $m_n, m_T, \ell_n > 1$ it results that $\mathbf{P}(\cdot), \mathbf{Q}(\cdot, \cdot), \mathbf{J}_\eta(\cdot)$ are continuously differentiable in a sufficiently small neighborhood of the equilibrium position.

In a sufficiently small neighborhood of the equilibrium position $(\mathbf{u}_0, \mathbf{0})$ the differential inclusion reduces to the equation

$$\mathbf{M}\ddot{\mathbf{u}} + \mathbf{C}\dot{\mathbf{u}} + \mathbf{K}\mathbf{u} + \mathbf{P}(\mathbf{u}) + \mathbf{Q}(\mathbf{u}, \dot{\mathbf{u}}) + \mathbf{J}_\eta(\mathbf{u}) = \mathbf{F}. \quad (5.8)$$

With $m_n, m_T, \ell_n > 1$ and taking $\mathbf{w}(t) = \mathbf{u}(t) - \mathbf{u}_0$ we obtain the linearized equation

$$\mathbf{M}\ddot{\mathbf{w}} + \mathbf{C}_0\dot{\mathbf{w}} + \mathbf{K}_0\mathbf{w} = \mathbf{0} \quad (5.9)$$

where

$$\mathbf{C}_0 \stackrel{\text{def}}{=} \mathbf{C} + \mathbf{C}^Q(\mathbf{u}_0) \quad (5.10)$$

$$\mathbf{K}_0 \stackrel{\text{def}}{=} \mathbf{K} + \mathbf{K}^P(\mathbf{u}_0) + \mathbf{K}^Q(\mathbf{u}_0, \mathbf{0}) + \mathbf{K}_\eta^J(\mathbf{u}_0) \quad (5.11)$$

and, for (\mathbf{w}, \mathbf{v}) in a small neighborhood of $(\mathbf{u}_0, \mathbf{0}) \in \mathbf{R}^3 \times \mathbf{R}^3$,

$$K^P(w) \stackrel{\text{def}}{=} \frac{\partial}{\partial w} P(w) \quad (5.12)$$

$$\begin{aligned} K^Q(w, v) &\stackrel{\text{def}}{=} \frac{\partial}{\partial w} Q(w, v) & C^Q(w) &= \frac{\partial}{\partial v} Q(w, v) \\ &= 0, \text{ if } v = 0 \end{aligned} \quad (5.13)$$

$$\begin{aligned} K_\eta^J(w) &\stackrel{\text{def}}{=} \frac{\partial}{\partial w} J_\eta(w) \\ &= \frac{\partial}{\partial w} J(w, v)|_{v=-\Phi} \end{aligned} \quad (5.14)$$

Linear stability analysis of the steady sliding equilibrium positions thus leads to the eigenvalue problem: find $\lambda \in \mathbb{C}$, $W \in \mathbb{C}^3$, $W \neq 0$ such that

$$[K_0 + \lambda C_0 + \lambda^2 M]W = 0. \quad (5.15)$$

5.2 Steady-Sliding Equilibrium and Linear Stability Analysis

5.2.1 Nondimensional form of the equations

The following nondimensional displacements will be used in most of the future developments in Section 5:

$$\tilde{u}_x = u_x/X, \tilde{u}_y = u_y/Y, \tilde{u}_\Theta = u_\Theta L/Y, \quad (5.16)$$

with

$$X = W/K_x, \quad (5.17)$$

$$Y = (W/c_n BL)^{1/m_n}. \quad (5.18)$$

For simplicity of notation, the (\sim) over the nondimensional displacement components (5.16) will be omitted wherever confusion is not likely to arise.

Assuming, for definiteness, $\dot{U}_x^C > 0$ ($\eta = +1$), the nondimensional form of the equilibrium equations (5.6) is thus:

$$u_{x0} = f \cdot I_x^J(u_{y0}, u_{\Theta 0}) \quad (5.19)$$

$$I_y^P(u_{y0}, u_{\Theta 0}) = 1 \quad (5.20)$$

$$I_{\Theta}^P(u_{y0}, u_{\Theta 0}) - \frac{f \cdot h}{2} I_x^J(u_{y0}, u_{\Theta 0}) = 0 \quad (5.21)$$

where we introduced the nondimensional parameters:

$$f = \frac{c_T}{c_n^{m_T/m_n}} \left(\frac{W}{BL} \right)^{\frac{m_T}{m_n} - 1} \quad (5.22)$$

$$h = \frac{H}{L} \quad (5.23)$$

and the functions

$$(w_y, w_{\Theta}) \in \mathbf{R}^2 \rightarrow I_x^J(w_y, w_{\Theta}) = \int_{-\frac{1}{2}}^{+\frac{1}{2}} [(w_y - \xi w_{\Theta})_+]^{m_T} d\xi$$

$$(w_y, w_{\Theta}) \in \mathbf{R}^2 \rightarrow I_y^P(w_y, w_{\Theta}) = \int_{-\frac{1}{2}}^{+\frac{1}{2}} [(w_y - \xi w_{\Theta})_+]^{m_n} d\xi$$

$$(w_y, w_{\Theta}) \in \mathbf{R}^2 \rightarrow I_{\Theta}^P(w_y, w_{\Theta}) = \int_{-\frac{1}{2}}^{+\frac{1}{2}} [(w_y - \xi w_{\Theta})_+]^{m_n} (-\xi) d\xi.$$

The parameters governing the steady-sliding equilibrium problem are thus the powers m_n and m_T , the friction parameter f and the geometric parameter h .

With respect to the eigenvalue problem (5.15), we observe that in the present case $K_{yx0} = K_{\Theta x0} = C_{yx0} = C_{\Theta x0} = M_{yx} = M_{\Theta x} = 0$, so that the characteristic equation for (5.15) decouples into

$$K_x + \lambda_x C_x + \lambda_x^2 M = 0 \quad (5.24)$$

$$\det(\mathbf{K}_0^* + \lambda^* \mathbf{C}_0^* + \lambda^{*2} \mathbf{M}^*) = 0. \quad (5.25)$$

Here, the superscript $*$ on a matrix denotes the submatrix associated with the normal and rotational degrees-of-freedom.

From (5.24) it is clear that the eigenvalues associated with the tangential motion are always imaginary ($C_x = 0$) or have negative real parts ($C_x > 0$).

In order to study the stability of normal and rotational motions, it is convenient to use the nondimensional normal and rotational displacements in (5.16) and the nondimensional eigenvalues

$$\Lambda = \lambda^*/\omega_{y0} \quad (5.26)$$

where ω_{y0} denotes the frequency of the free normal oscillation of the block for the linearized normal stiffness at the (frictionless) equilibrium position $u_{y0} = 1$, i.e.,

$$\omega_{y0} = (m_n c_n BLY^{m_n-1}/M)^{1/2}. \quad (5.27)$$

The characteristic equation (5.25) becomes then

$$\begin{aligned} & \det \left(\int_{-\frac{1}{2}}^{+\frac{1}{2}} [a_0(\xi)]^{m_n-1} \begin{bmatrix} 1 & -\xi \\ -\xi & \xi^2 \end{bmatrix} d\xi - \frac{f m_T h}{2 m_n} \int_{-\frac{1}{2}}^{+\frac{1}{2}} [a_0(\xi)]^{m_T-1} \begin{bmatrix} 0 & 0 \\ 1 & -\xi \end{bmatrix} d\xi \right. \\ & \left. + 2\Lambda \hat{z} \int_{-\frac{1}{2}}^{+\frac{1}{2}} [a_0(\xi)]^{\ell_n} \begin{bmatrix} 1 & -\xi \\ -\xi & \xi^2 \end{bmatrix} d\xi + \Lambda^2 \begin{bmatrix} 1 & 0 \\ 0 & (1+h^2)/12 \end{bmatrix} \right) = 0 \end{aligned} \quad (5.28)$$

where we have used the notation $a_0(\xi) = (u_{y0} - \xi u_{\theta 0})_+$ and we have introduced the interface normal damping parameter

$$\hat{z} = \frac{\sqrt{W/M}}{2\sqrt{m_n}} \frac{b_n}{c_n^{(\ell_n+1/2)/m_n}} \left(\frac{W}{BL} \right)^{(\ell_n+1/2)/m_n-1}$$

It is clear from (5.28) that the parameters that govern the linear stability of the steady-sliding are the powers m_n, m_T and ℓ_n , the friction parameter f , the normal damping parameter \hat{z} and the geometric parameter h . It is also clear that the contribution of the frictional resistance to the eigenvalue problem (5.28) is a nonsymmetric matrix.

5.2.2 Numerical results

In order to study numerically the steady-sliding equilibrium and its linear stability we first select values for m_n, m_T and h and solve the problem (5.19, 5.20, 5.21) for increasing values of f in the range $[0, \bar{f})$, where $\bar{f} = \bar{f}(m_n, m_T, h)$ denotes the value of f for which steady sliding equilibrium ceases to be possible due to the tumbling of the block (for $m_n = m_T$ it is easy to see that $\bar{f} = 1/h = L/H$). For each of the values of f considered and for some ℓ_n and \hat{z} the eigenvalue problem (5.15) is then solved numerically. Typical results for the normal (u_{y0}) and rotational ($u_{\theta 0}$) equilibrium displacements as the friction parameter is increased

are shown in Figs. 5.2 and 5.3. It is clear that, as should be expected, both u_{y0} and $u_{\theta 0}$ grow unboundedly as f approaches \bar{f} . In Figs. 5.4 and 5.5 we plot typical evolutions of the eigenvalues Λ as the friction parameter f is increased, when the normal interface damping is zero (Fig. 5.4) or different from zero (Fig. 5.5).

It is clear that, in both cases, *for some range of f , the nonsymmetry of the friction contributions in (5.28) originates eigenvalues Λ with positive real part, and this implies the dynamic instability of the corresponding steady sliding equilibrium positions.*

These regions of instability are identified in the (h, f) -parameter plane (Figs. 5.6, 5.7 and 5.8) for various values of m_n, m_T, ℓ_n and \hat{z} . For each $h > 0$, those regions of instability are delimited from below by the values (f_1) at which a pair of conjugate eigenvalues Λ starts to have positive real parts and from above by the values (f_2) at which such positive real parts cease to exist.

Since our numerical results were obtained by incrementing the value of f without searching for the exact values of f_1 and f_2 , these values are represented in Figs. 5.6 and 5.7 by the pairs of endpoints of the increments of f at which the above mentioned transitions occur. For clarity, in Fig. 5.8, the transitional increments' endpoints lying inside the instability region are omitted. When $m_T \neq m_n$ the friction parameter f does not coincide with what may be called the *steady-sliding coefficient of friction*, i.e., the ratio between the steady-sliding friction and normal forces (recall 5.19 and 5.20):

$$\mu_0 = u_{x0} = f \cdot I_x^J(u_{y0}, u_{\theta 0})$$

For this reason in Fig. 5.7 ($m_n = 2, m_T = 3$) the boundaries of the instability region are plotted in terms of both f and μ_0 . Finally we note that similar computations were also done for a case ($m_n = 3, m_T = 2$) that corresponds to a decrease of the friction parameter with the increase of the average normal pressure according to $f = (c_T/c_n^{2/3})(W/BL)^{-1/3}$. In this case, in the absence of normal interface damping, the eigenvalues of (5.28) were imaginary for all the values of the parameters h and f employed.

We also remark that, if some eigenvalue of (5.15) has a zero real part, no definitive conclusion on the stability of the equilibrium solution can be obtained from that eigenvalue problem. Since we are mostly interested in showing that, for some range of the parameters involved, steady-sliding is unstable, we shall not pursue here the study of what happens outside the region of instability in Figs. 5.6, 5.7, and 5.8 when $\hat{z} = 0$. However, *if $C_x > 0$ and $b_n > 0$ ($\hat{z} > 0$), the admissible points outside the region of instability in Fig. 5.6 correspond to asymptotically stable steady-sliding positions.*

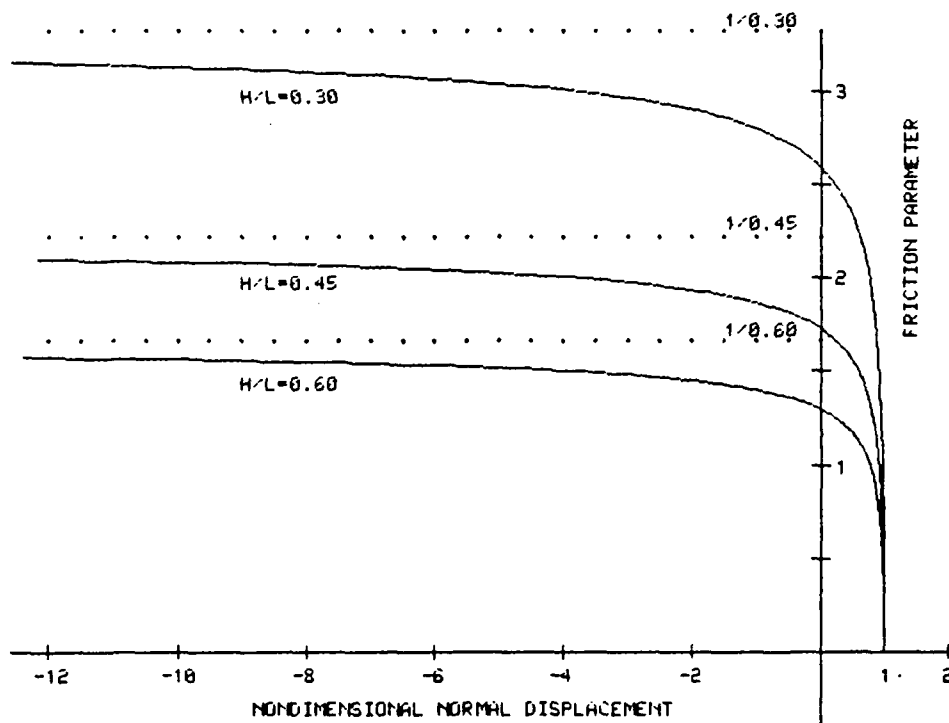


Figure 5.2: Steady-sliding normal displacement (u_{y0}) of a rigid block as a function of the friction parameter ($f \equiv \mu$) for $m_T = m_n = 2.5$ and three values of the geometric parameter ($h = H/L$).

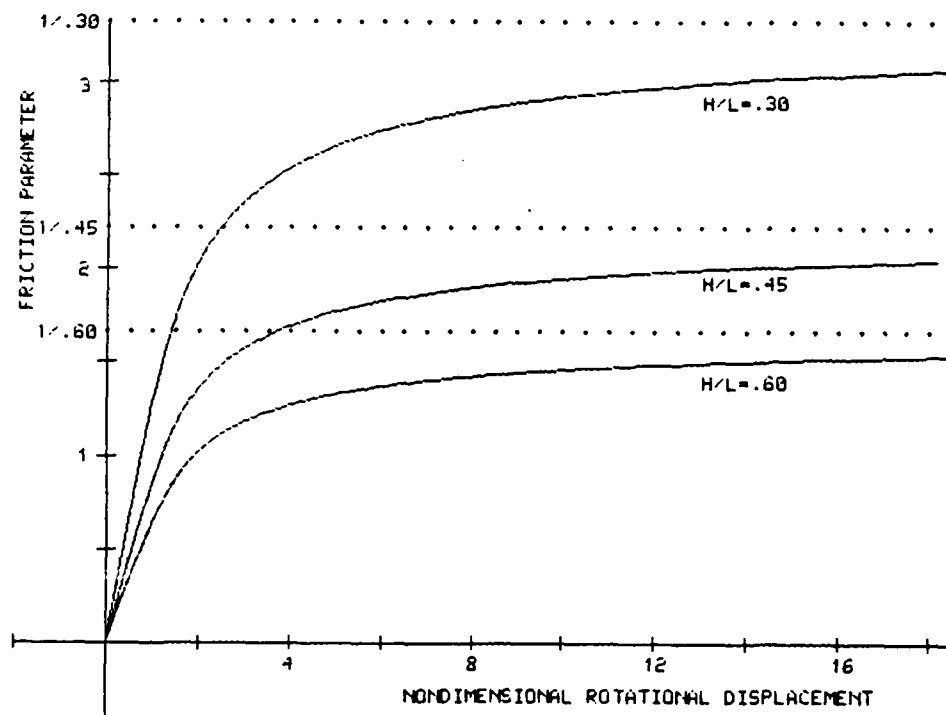


Figure 5.3: Steady-sliding rotational displacement ($u_{\theta 0}$) of a rigid block as a function of the friction parameter ($f \equiv \mu$) for $m_T = m_n = 2.5$ and three values of the geometric parameter ($h = H/L$).

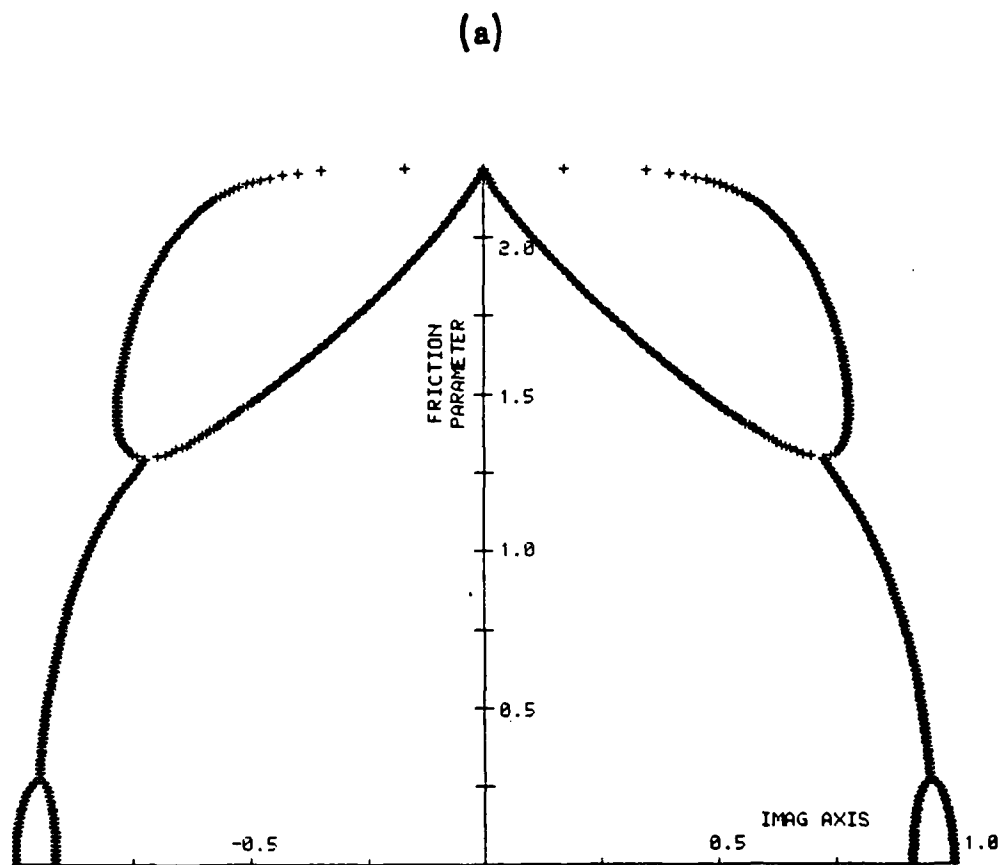
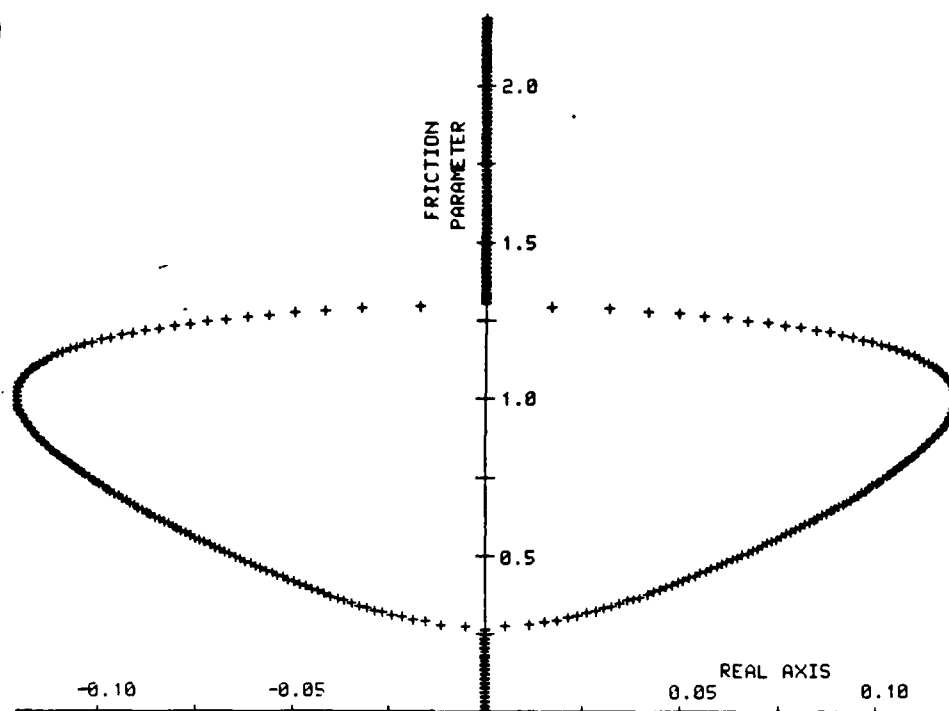
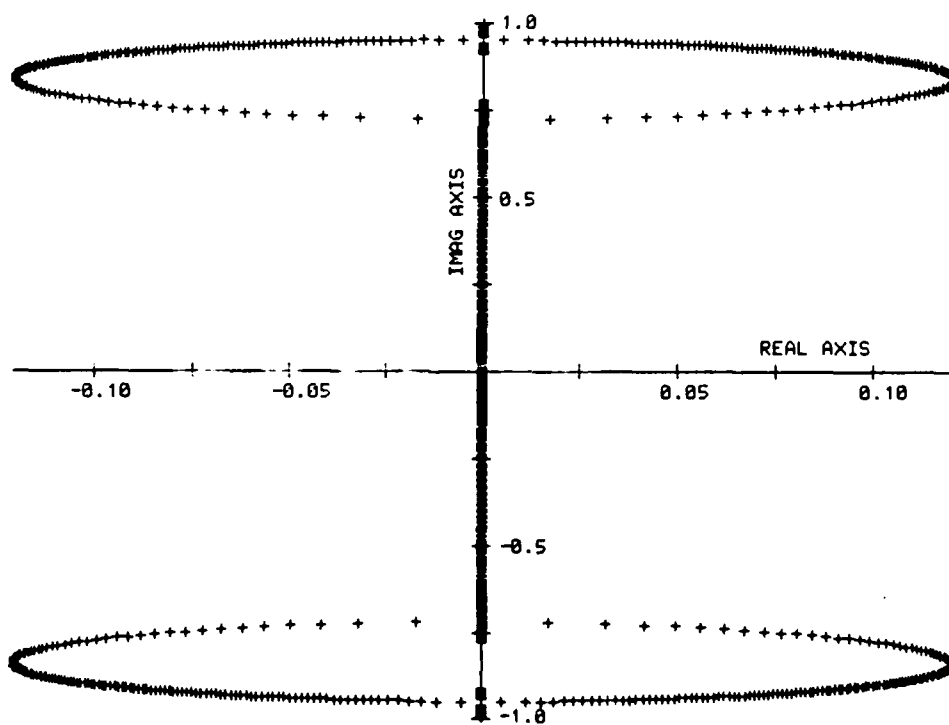


Figure 5.4: Part (a). Orthographic projections (a,b,c) and perspective (d) of the root curves of the characteristic equation (5.28) for the admissible range of the friction parameter $f \in [0, 1/0.45)$ with $m_n = m_T = 2.5$, $h = 0.45$ and $\hat{z}=0.00$ fixed.

(b)



(c)



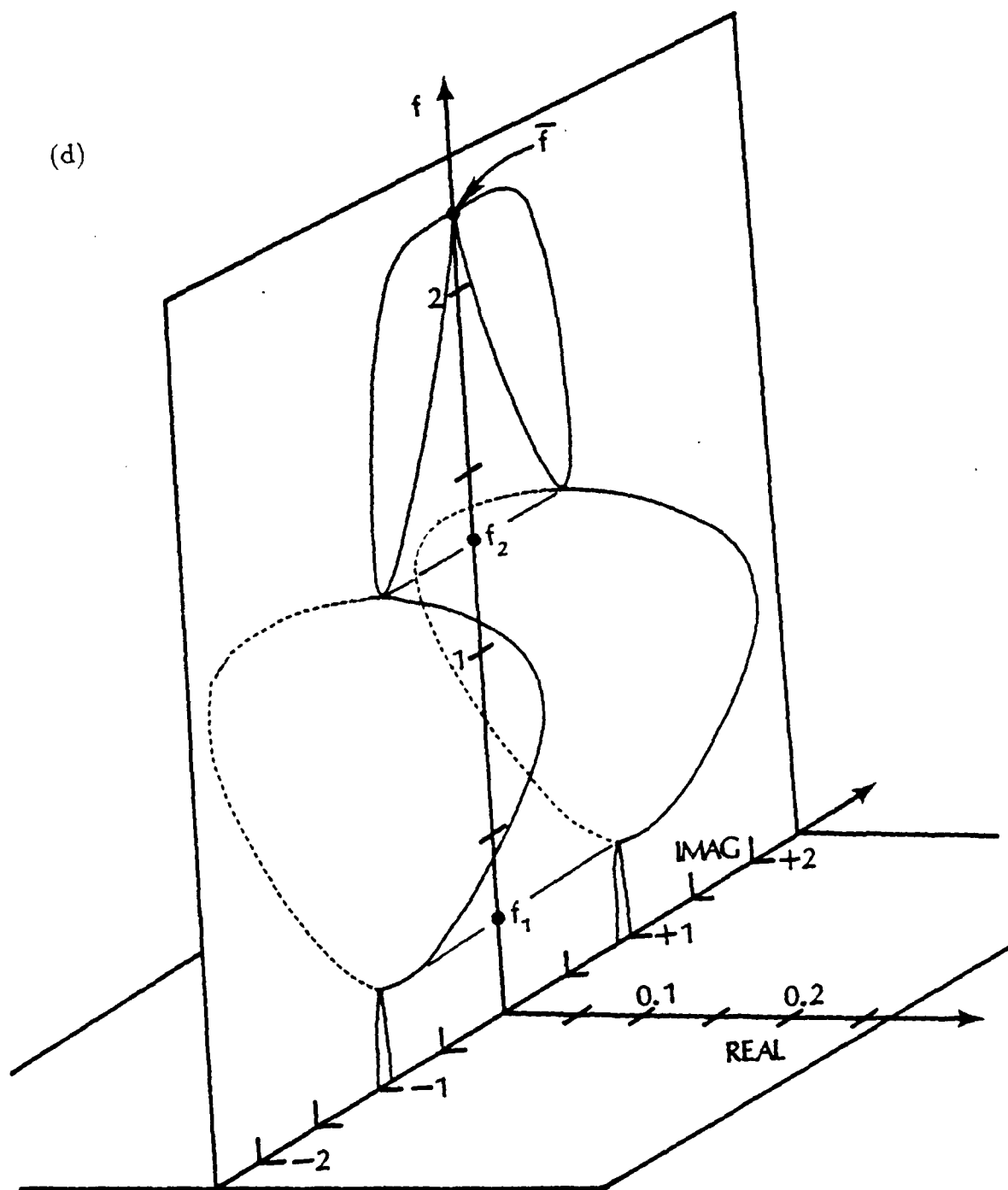


Figure 5.4: Part (d)

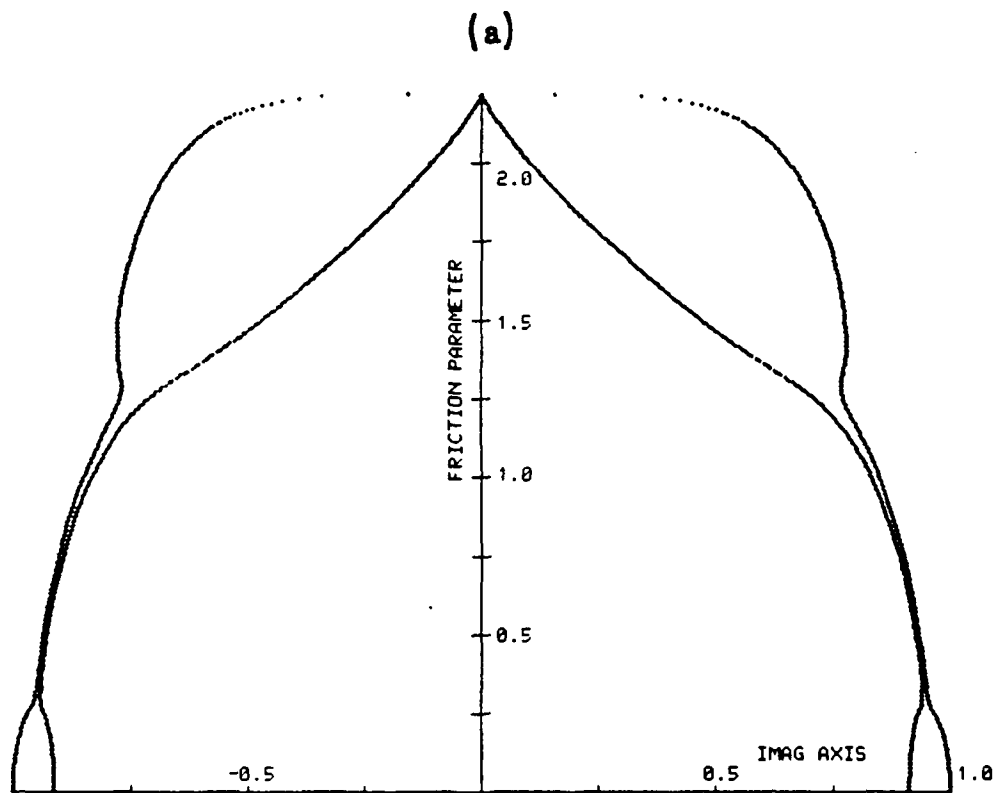
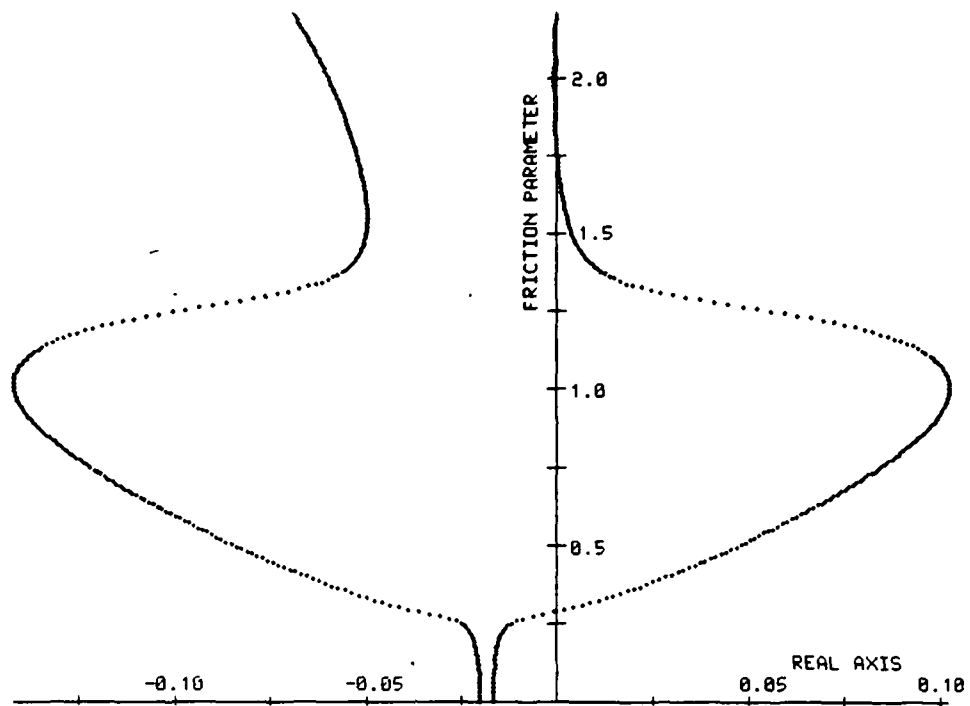


Figure 5.5: Part (a). Orthographic projections of the root curves of the characteristic equation (5.28) for the admissible range of the friction parameter $f \in [0, 1/0.45)$ with $m_n = m_T = \ell_n = 2.5$, $h = 0.45$ and $\hat{z} = 0.02$ fixed.

(b)



(c)

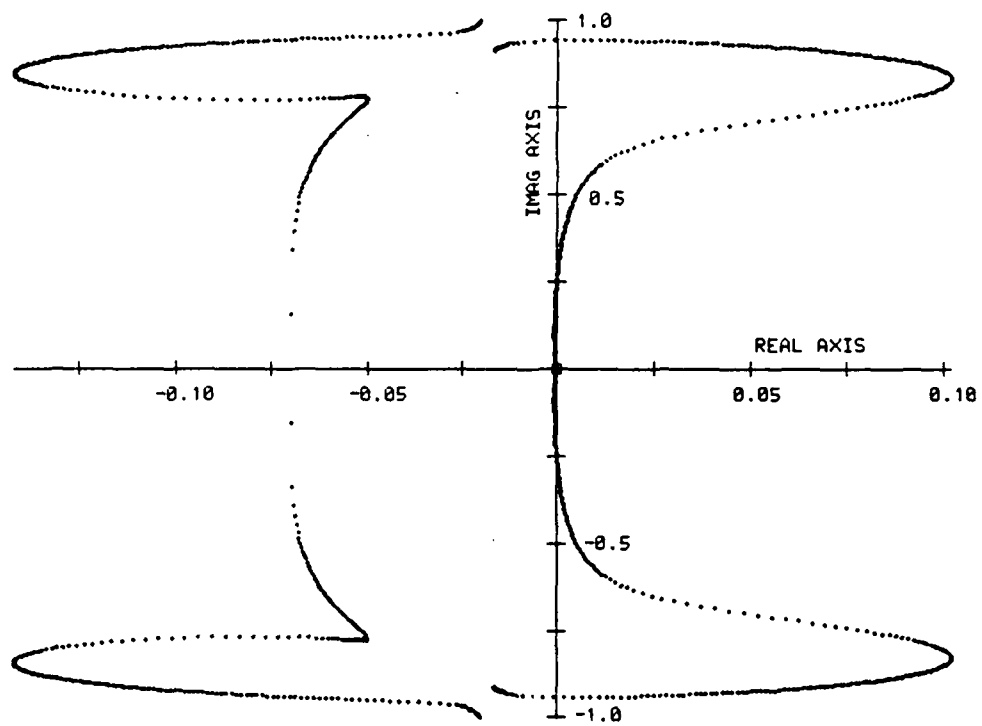


Figure 5.5: Parts (b) and (c)

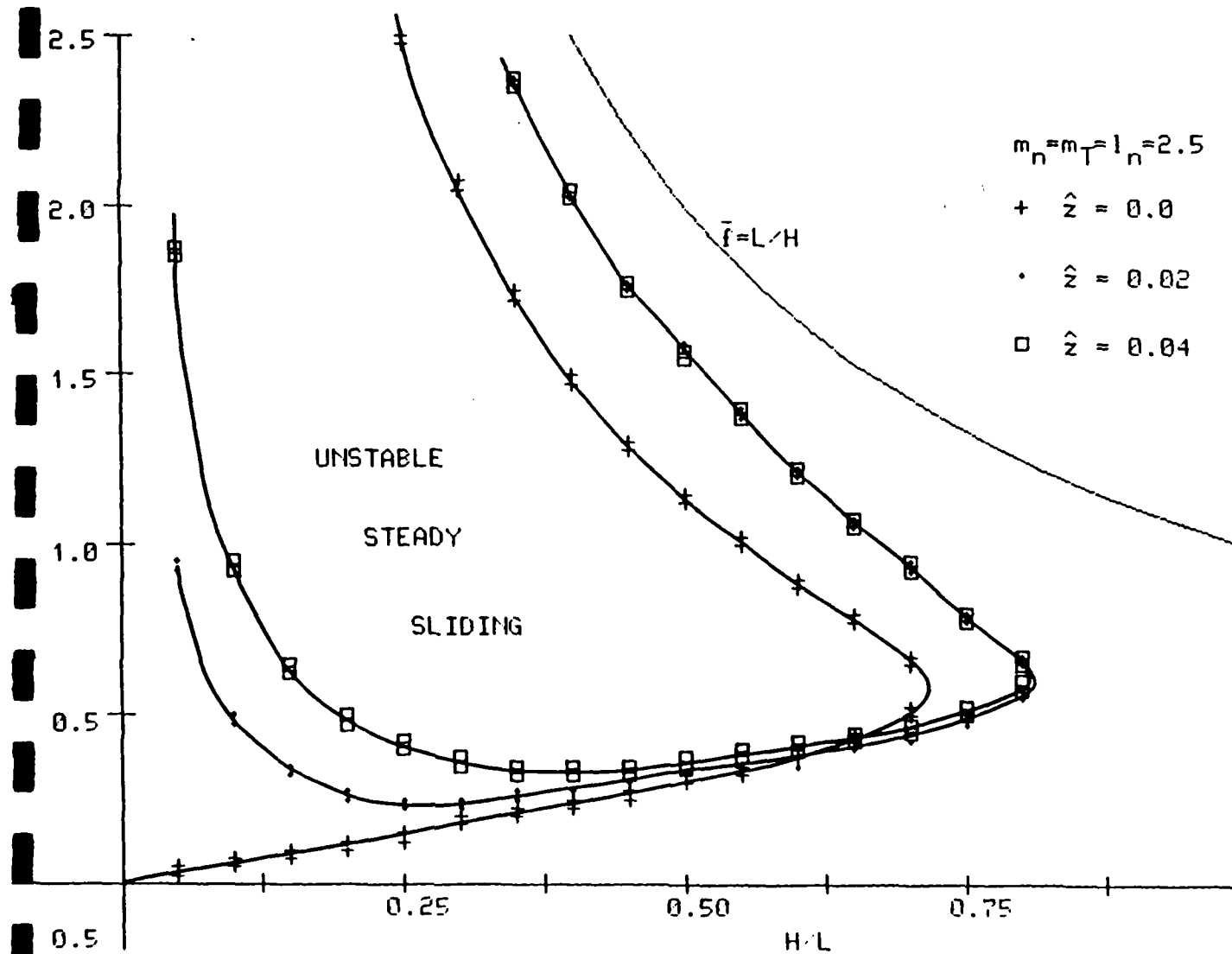


Figure 5.6: Effect of the normal interface damping on the stability of the steady-sliding equilibrium of a rigid block.

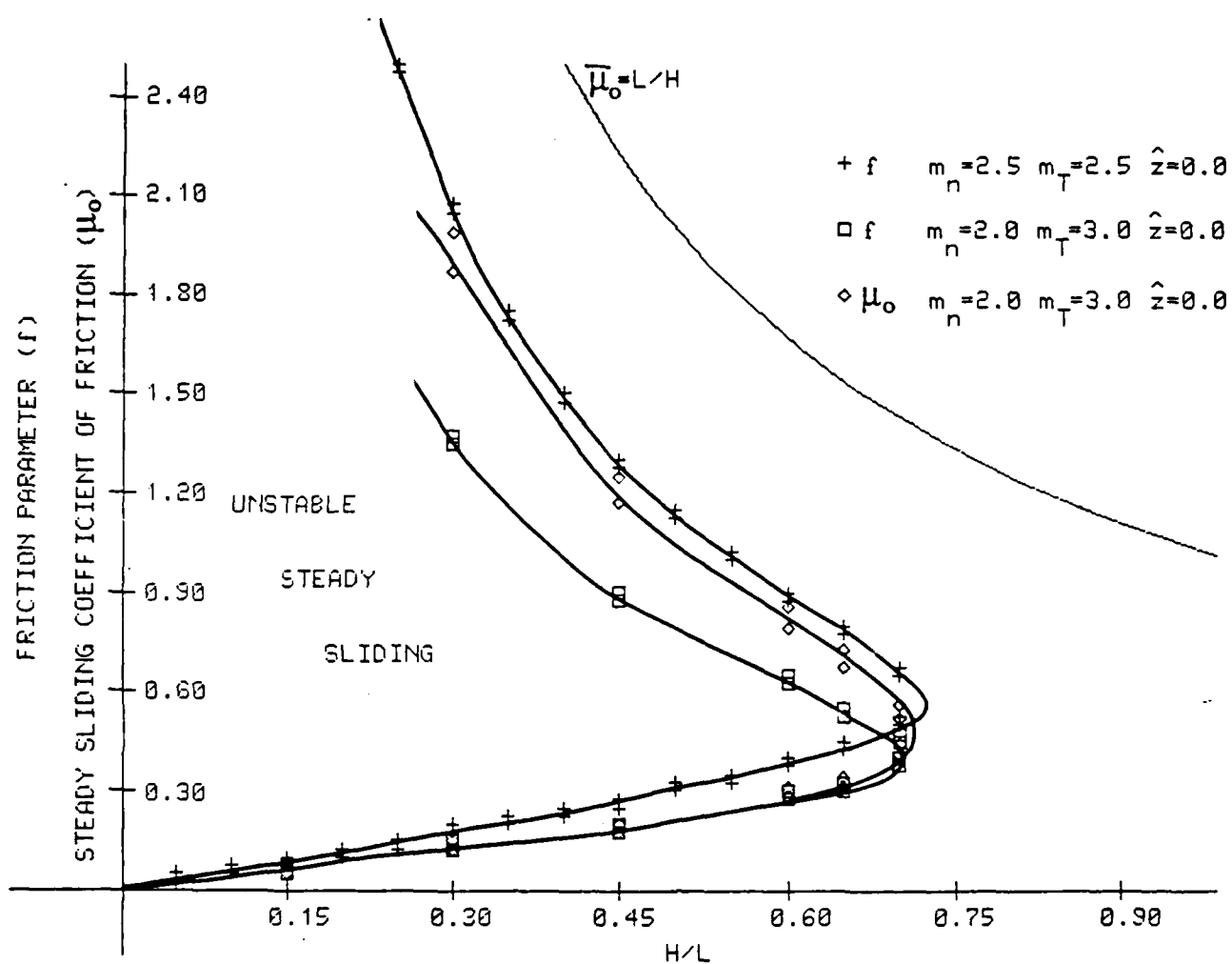


Figure 5.7: The regions of instability for $m_n = m_T = 2.5$ and for $m_n = 2$, $m_T = 3.0$, with $\hat{z} = 0.0$.

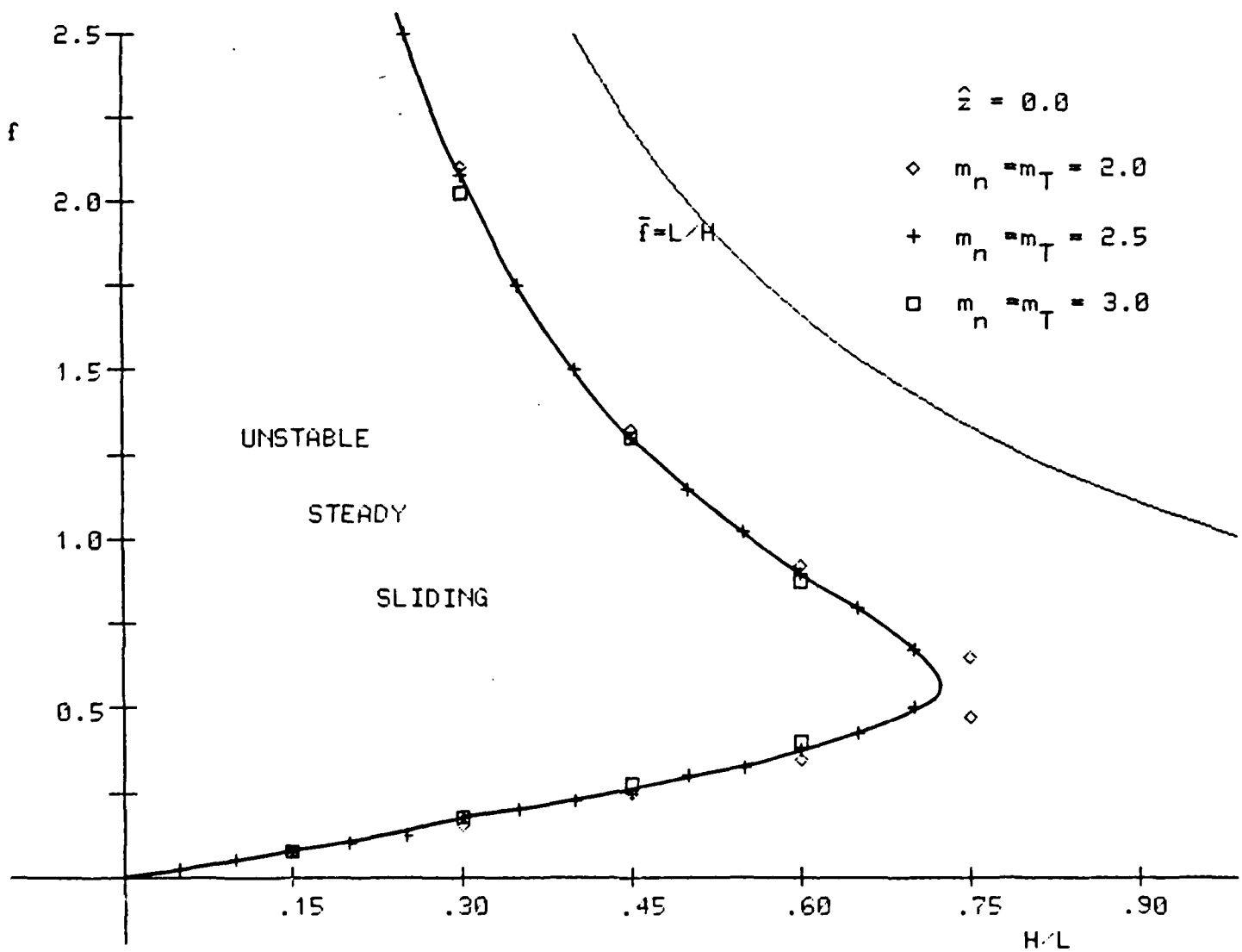


Figure 5.8: Effect of the value of the powers $m_n = m_T$ on the stability of the steady-sliding equilibrium of a rigid block.

5.2.3 Discussion

We observe that the small rotations assumption adopted in this work (recall Section 5.1) implies that, for f close to \bar{f} , the results obtained here do not represent the true behavior of the system, since, for such values of f , large rotations do occur. We remark that these limitations on the validity of our results do not affect the fundamental observation of the previous section, i.e., that for some range of f and appropriate values of the other parameters, steady-sliding is not stable: it can be checked in Figs. 5.3., 5.6, 5.7, and 5.8 that the transitions on the nature of the eigenvalues of (5.28) occur when the rotations $u_{\theta 0}$ are still very small (note that for L in the range 1 to 100 cm and Y in the range 0.3 to 10 μm the physical values in rad of the steady-sliding rotations are obtained from the nondimensional values in Fig. 5.3 through multiplication by a factor in the range 3×10^{-7} to 10^{-3}). Only somewhere above the upper boundary of instability in Figs. 5.6, 5.7, and 5.8, the small rotations assumption ceases to be valid. Furthermore, we believe that, in practice, the geometric nonlinearity does not play any significant role in block-on-slideway sliding systems of the type studied in this section: it is reasonable to expect that no one would operate or run an experiment with such a system allowing for the occurrence of large rotations. Of course, the same may not be true in other circumstances, namely with some pin-on-disk friction apparatus having very flexible arms and very small contact regions.

From Fig. 5.8 it can be concluded that, when the coefficient of friction is independent of the normal pressure, *the specific value of $m_n (= m_T)$ in the typical range [2,3] does not affect considerably the boundaries of the region of instability.* In the case $2 = m_n < m_T = 3$, if we consider the stability boundary in terms of the (measurable) μ_0 we again see that it is not much different from the cases with $m_n = m_T$. The absence of a region of linear instability for the case $3 = m_n > m_T = 2$ seems rather interesting and deserving further study. At this writing, we don't know of any experimental work correlating stability of steady sliding with normal pressure dependence of the coefficient of friction, which might provide a basis for a qualitative evaluation of these numerical results.

In Fig. 5.6 it can be observed that, for small values of the geometric parameter $h (= H/L)$, the interface normal damping increases the value of the friction parameter at which the instability initiates, i.e., it reduces the region of instability. However, it is also clear from the same figure that the normal damping moves up the upper boundary of the instability region, i.e., it increases the region of instability. Hence, addition of normal interface damping in the form adopted here may indeed have a destabilizing effect on the steady-sliding equilibrium for some ranges of values of the parameters involved. Situations of this type, in which

positive viscous "damping" contributions may have a destabilizing effect, are common in nonconservative systems. Such effects were first discovered by Ziegler [79] and have since been analyzed by various authors in connection with the study of critical loads for beams subjected to follower forces, stability of fluid conveying pipes, panel flutter, etc. (see, e.g., the books of Ziegler [80], Leipholz [53], Huseyin [41], Guckenheimer and Holmes [36]). A detailed discussion of these effects in the present problem falls outside the scope of this work. Here we only remark that, *with a deformable block and with no normal interface damping*, the steady-sliding equilibrium is unstable in the region of the (h, f) -parameter plane where the destabilizing effects of the normal damping are more significant (the region above the upper boundary of the *rigid body* instability); see Oden and Martins [56].

5.3 Low-Frequency Stick-Slip Motion and Apparent Reductions of Kinetic Friction

5.3.1 Nondimensional form of the equations of motion

We denote by τ the nondimensional time

$$\tau = \omega t \quad (5.29)$$

and we choose ω to be the frequency of the free tangential oscillation of the block, i.e.,

$$\omega = (K_x/M)^{1/2}. \quad (5.30)$$

Using again the nondimensional displacements defined in (5.16), (5.17) and (5.18) and omitting again the (\sim) over the nondimensional displacement components, the governing system (5.4) becomes

$$\begin{aligned} M\mathbf{u}''(\tau) + C\mathbf{u}'(\tau) + K\mathbf{u}(\tau) \\ + P(\mathbf{u}(\tau)) + Q(\mathbf{u}(\tau), \mathbf{u}'(\tau)) \\ + J(\mathbf{u}(\tau), \mathbf{u}'(\tau) - \Phi') \ni \mathbf{F} \end{aligned} \quad (5.31)$$

where, now,

$$M = \begin{bmatrix} 1 & 0 & 0 \\ 0 & s^2 & 0 \\ 0 & 0 & s^2(1+h^2)/12 \end{bmatrix}; K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; C = 2 \begin{bmatrix} z_x & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$P(\omega) = s \int_{-\frac{1}{2}}^{+\frac{1}{2}} [a(\xi)]^{m_n} \begin{Bmatrix} 0 \\ 1 \\ -\xi \end{Bmatrix} d\xi$$

$$Q(\omega, \nu) = 2\hat{z}(m_n)^{\frac{1}{2}} s^{3/2} \int_{-\frac{1}{2}}^{+\frac{1}{2}} [a(\xi)]^{\ell_n} (\nu_y - \xi \nu_\Theta) \begin{Bmatrix} 0 \\ 1 \\ -\xi \end{Bmatrix} d\xi \quad (5.32)$$

$$J(\omega, \nu) = f \operatorname{sgn}(\nu_x + \frac{sh}{2} \nu_\Theta) \begin{Bmatrix} 1 \\ 0 \\ \frac{sh}{2} \end{Bmatrix} \int_{-\frac{1}{2}}^{+\frac{1}{2}} [a(\xi)]^{m_\tau} d\xi \quad (5.33)$$

$$\Phi' = \begin{Bmatrix} U_x^C \\ 0 \\ 0 \end{Bmatrix}, U_x^C = \dot{U}_x^C / \omega X \quad (5.34)$$

$$F = \begin{Bmatrix} 0 \\ s \\ 0 \end{Bmatrix}. \quad (5.35)$$

Here $a(\xi) = (\omega_y - \xi \omega_\Theta)_+$, $()'$ denotes again differentiation with respect to the nondimensional time τ , and, in addition to the parameters f, h and \hat{z} defined earlier, we have introduced the parameters

$$s = Y/X = \frac{K_x}{W} \left(\frac{W}{c_n BL} \right)^{1/m_n} \quad (5.36)$$

$$z_x = \frac{C_x}{2\sqrt{MK_x}} \quad (5.37)$$

The *stiffness parameter* s has the same meaning as in Section 4.1. Note that when $s = m_n$ the tangential stiffness K_x is equal to the (linearized) normal stiffness $K_{on} = m_n c_n B L Y^{m_n-1}$ at the (nondimensional) frictionless equilibrium position $u_{y0} = 1$. The parameter z_x is the (usual) *tangential damping parameter*. The initial conditions (5.5) become now

$$u(0) = \bar{u}_0, u'(0) = \bar{u}_1 \quad (5.38)$$

Finally we observe that, for computational purposes, problem (5.31), (5.38) is regularized using the procedure employed in Section 4.1. $\Delta\tau_{\max}$ denotes the nondimensional maximum time step for the numerical integration of the equations of motion (see Oden and Martins [56]).

5.3.2 Numerical results and discussion

A complete qualitative study of the system (5.31), (5.38) is not available yet. Here we present several numerical studies designed to reveal the effect that some of the governing parameters have on the behavior of the system and, whenever possible, we qualitatively compare our numerical results with experimental observations.

The physically interesting situations arise when the data is chosen such that the steady-sliding equilibrium is unstable. In the examples studied in the remainder of this section the following common data is used

$$\begin{aligned} m_n = m_T = l_n &= 2.5 \\ h &= 0.45 \\ f &= 0.6 \end{aligned} \tag{5.39}$$

Low frequency stick-slip motion and apparently smooth sliding. In the computations described next we used, in addition, the following data

$$s = 0.01 \tag{5.40}$$

$$z_x = 0.001 \tag{5.41}$$

$$\hat{z} = 0.01 \tag{5.42}$$

$$\epsilon/U_x'^C \leq 0.1 \tag{5.43}$$

$$\Delta\tau_{max} \leq \frac{2\pi}{100} \frac{\omega}{\omega_{y0}} \tag{5.44}$$

and various values of the driving velocity $U_x'^C$ as indicated on the figures. The initial conditions were the following

$$\bar{u}_{x0} = u_{x0}, \bar{u}_{y0} = u_{y0} + {}^p u_y, \bar{u}_{\theta 0} = u_{\theta 0}, \bar{\mathbf{u}}_1 = \mathbf{0}, \tag{5.45}$$

where ${}^p u_y$ is a small normal displacement perturbation:

$$^p u_y = -0.01 \text{ or } -0.001. \quad (5.46)$$

The values of s and z_x considered above are "in the small range" and the initial conditions correspond to a small normal perturbation of the steady-sliding equilibrium state.

The following remarks give a summary description and interpretation of the numerical results obtained for the conditions indicated above:

1. Due to the instability of the normal and rotational modes, the normal and rotational oscillations grow (see Fig. 5.9).
2. The variation of the normal force on the contact produces changes in the sliding friction force which in turn produce a tangential oscillation.
3. The tangential oscillation may then become sufficiently large that, for small values of the driving velocity U_x^C , the points of the body on the the contact surface attain the velocity U_x^C and the body sticks for short intervals of time (see Fig. 5.10).
4. With the growing of the normal oscillation actual normal jumps of the body may occur (see Fig. 5.9).
5. The repeated periods of adhesion have the result of decreasing the average value of the friction force on the contact and, due to the absence of equilibrium with the restoring force on the tangential spring, the tangential displacement of the center of mass decreases (see Figs. 5.11, and 5.12).
6. Then, one of the two following situations may occur:
 - (a) *for values of U_x^C larger than some critical value*, the normal, rotational and tangential oscillations evolve to what appears to be a steady oscillation with successive periods of adhesion and sliding, the average values of the friction force and of the spring elongation being smaller than those corresponding to the steady-sliding equilibrium position (see Figs. 5.11 and 5.13 to 5.16).
 - (b) *for values of U_x^C lower than the critical value*, and at a sufficiently small value of the spring elongation, the normal interface damping is able to damp out the normal (and rotational) oscillation (see Fig. 5.17) and the body sticks (see Fig. 5.12) since the restoring force of the spring is then smaller than the maximum available friction force.

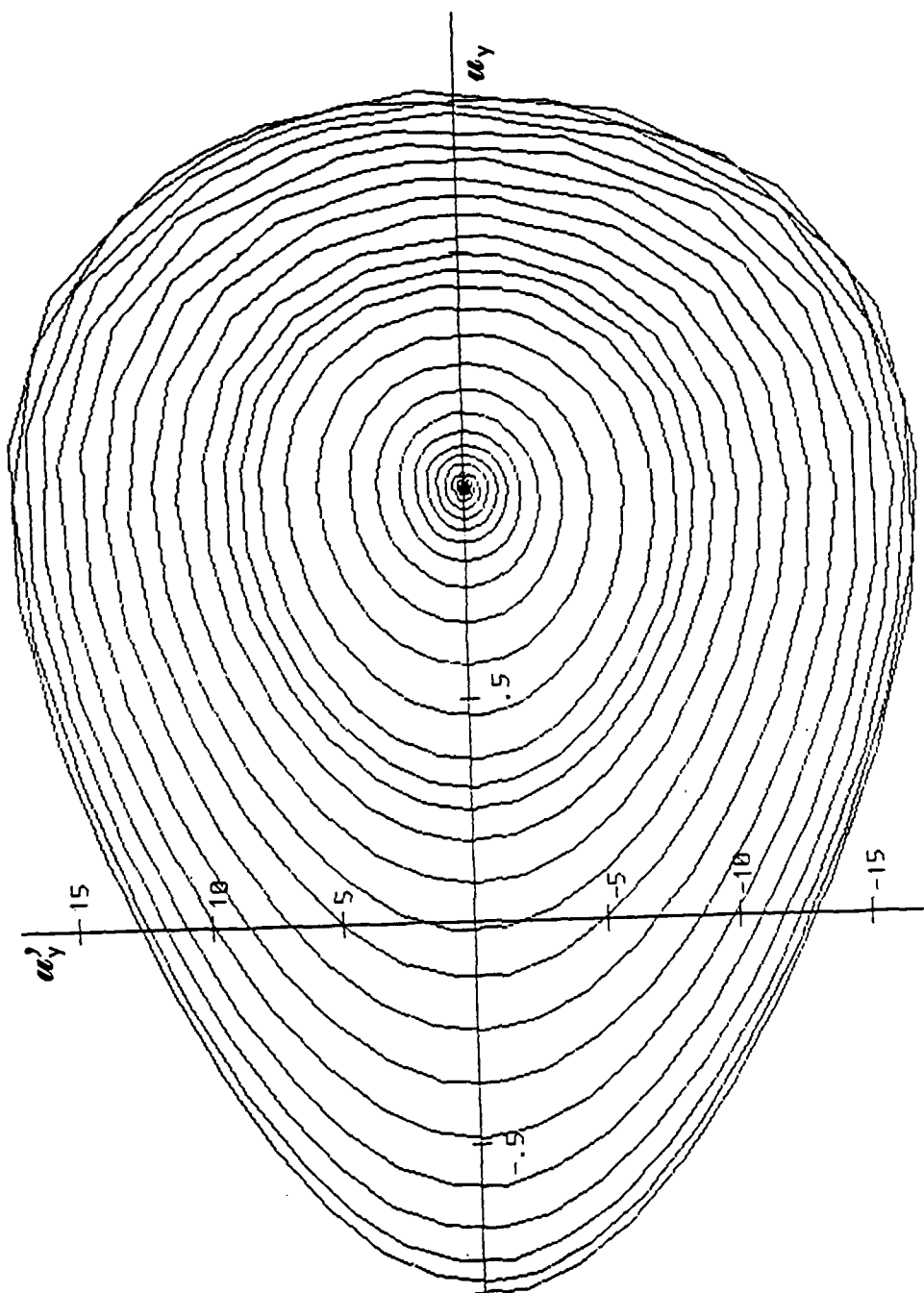


Figure 5.9: Phase plane plot of the growing portion of the normal oscillation of the center of mass in the course of low-frequency stick-slip motion. $U_x^C = 0.01$. (Note: not all the computed points are plotted).

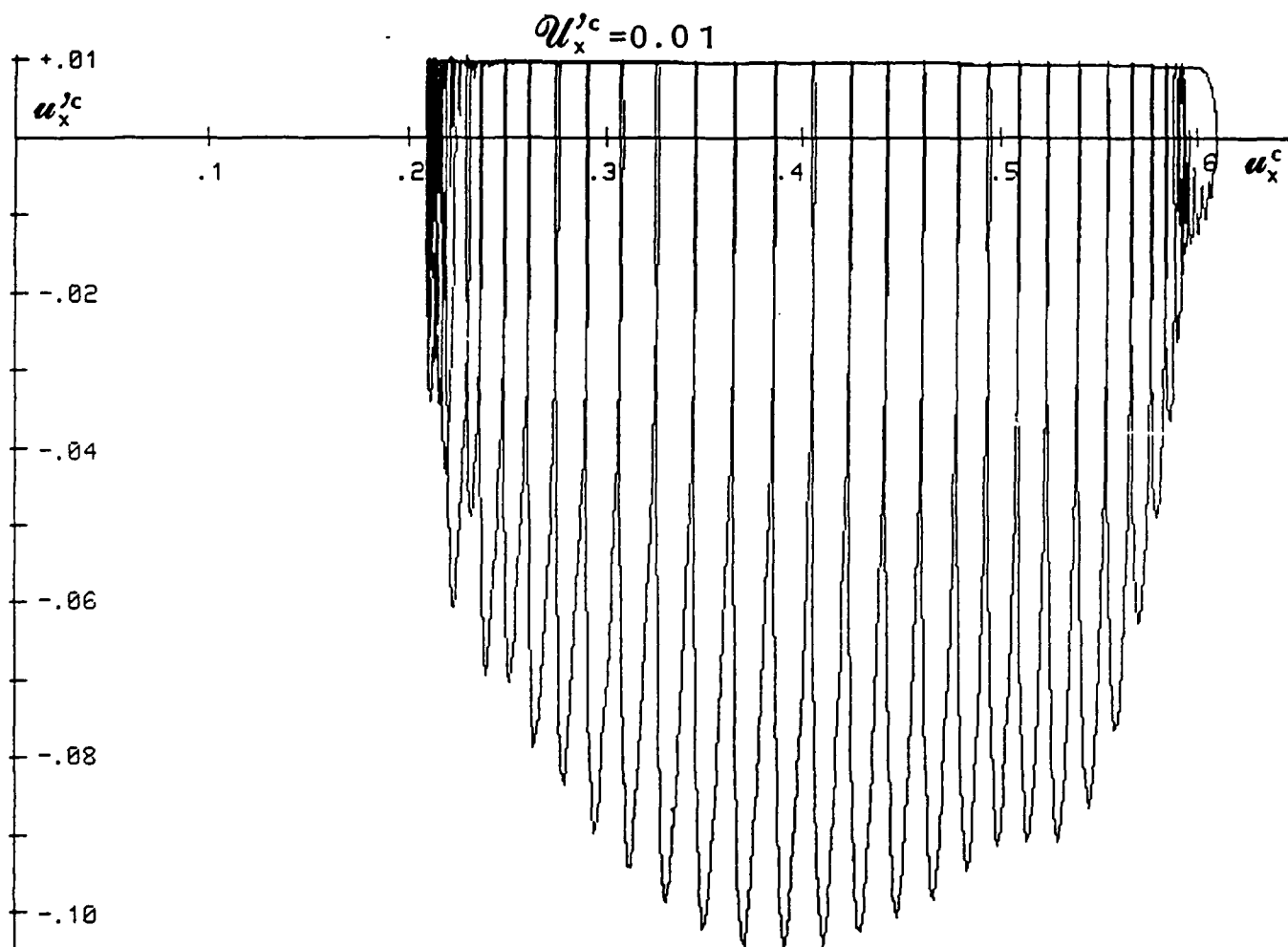


Figure 5.10: Phase plane plot of the tangential motion of the points of the block on the contact surface during one period of the low-frequency stick-slip motion.

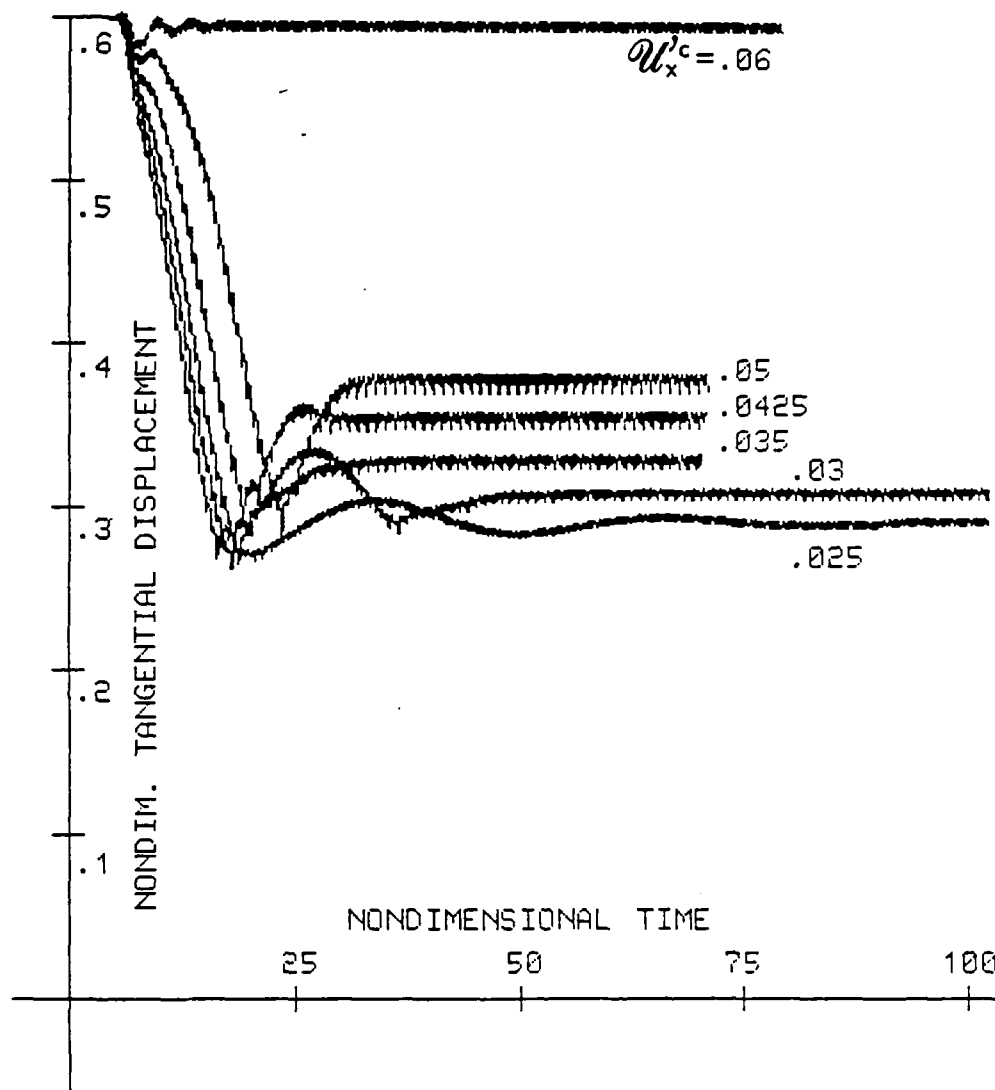


Figure 5.11: Apparently smooth sliding for small s , small z_x , large U_x^C ($s=0.01$, $z_x = 0.001$ and U_x^C successively equal to 0.025, 0.030, 0.035, 0.0425, 0.05, 0.06).

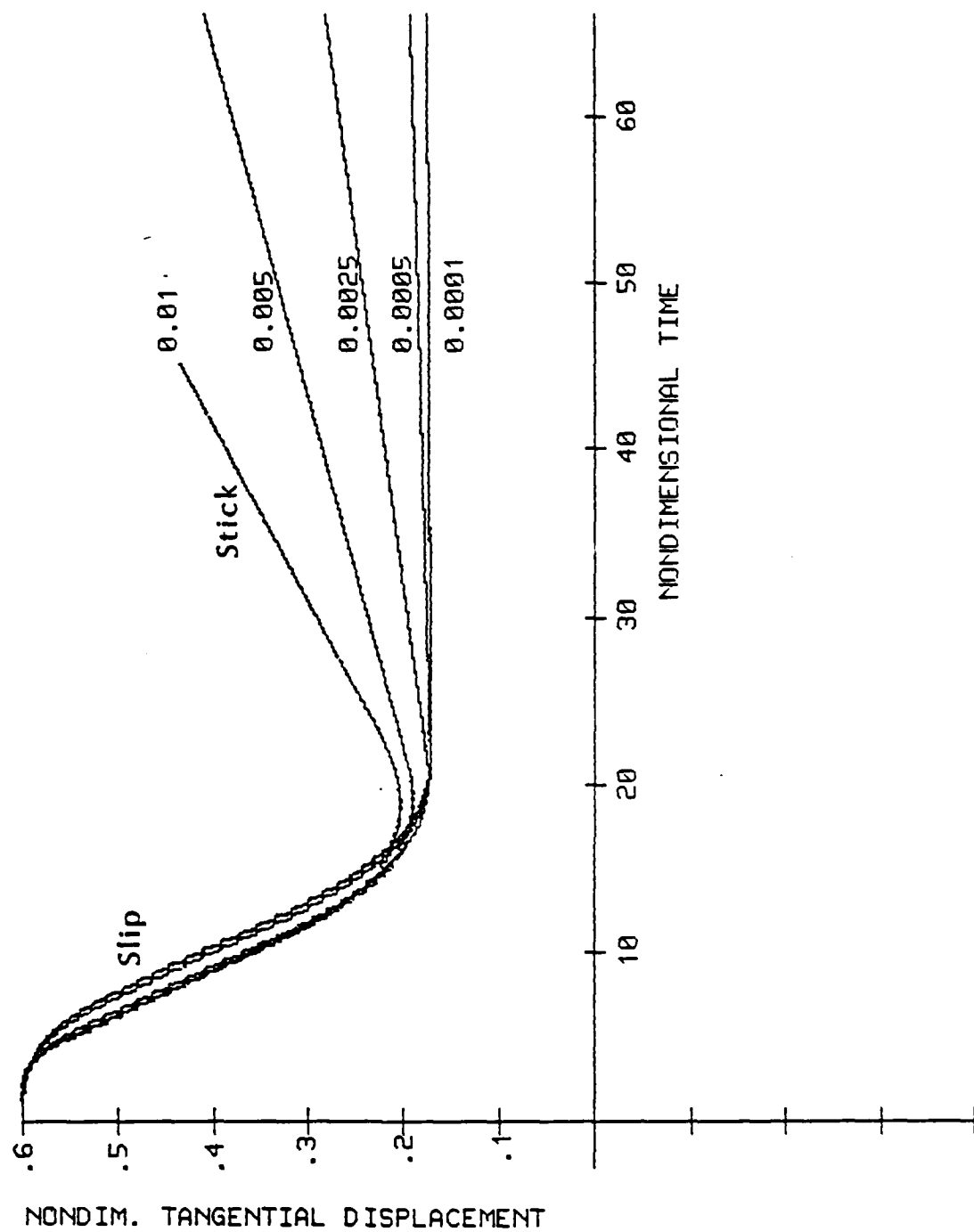


Figure 5.12: Low-frequency stick-slip motions for various small driving velocities ($s=0.01$, $z_x=0.001$ and U_x^C successively equal to 0.01, 0.005, 0.0025, 0.0005, 0.0001).

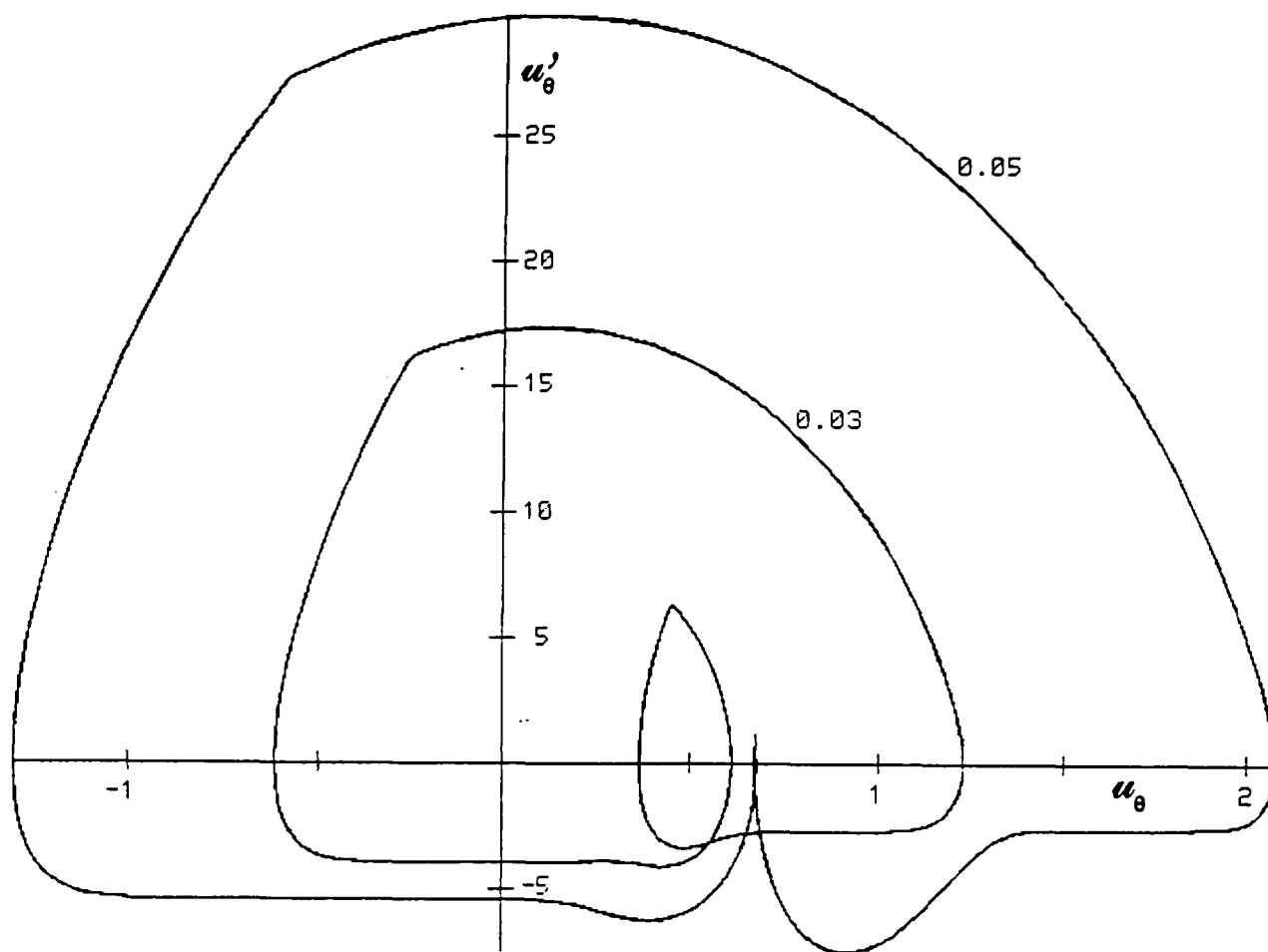


Figure 5.13: Phase plane plots for the rotation degree-of-freedom ($s=0.01$, $z_x=0.001$ and $U_x^C=0.03$ and 0.05).

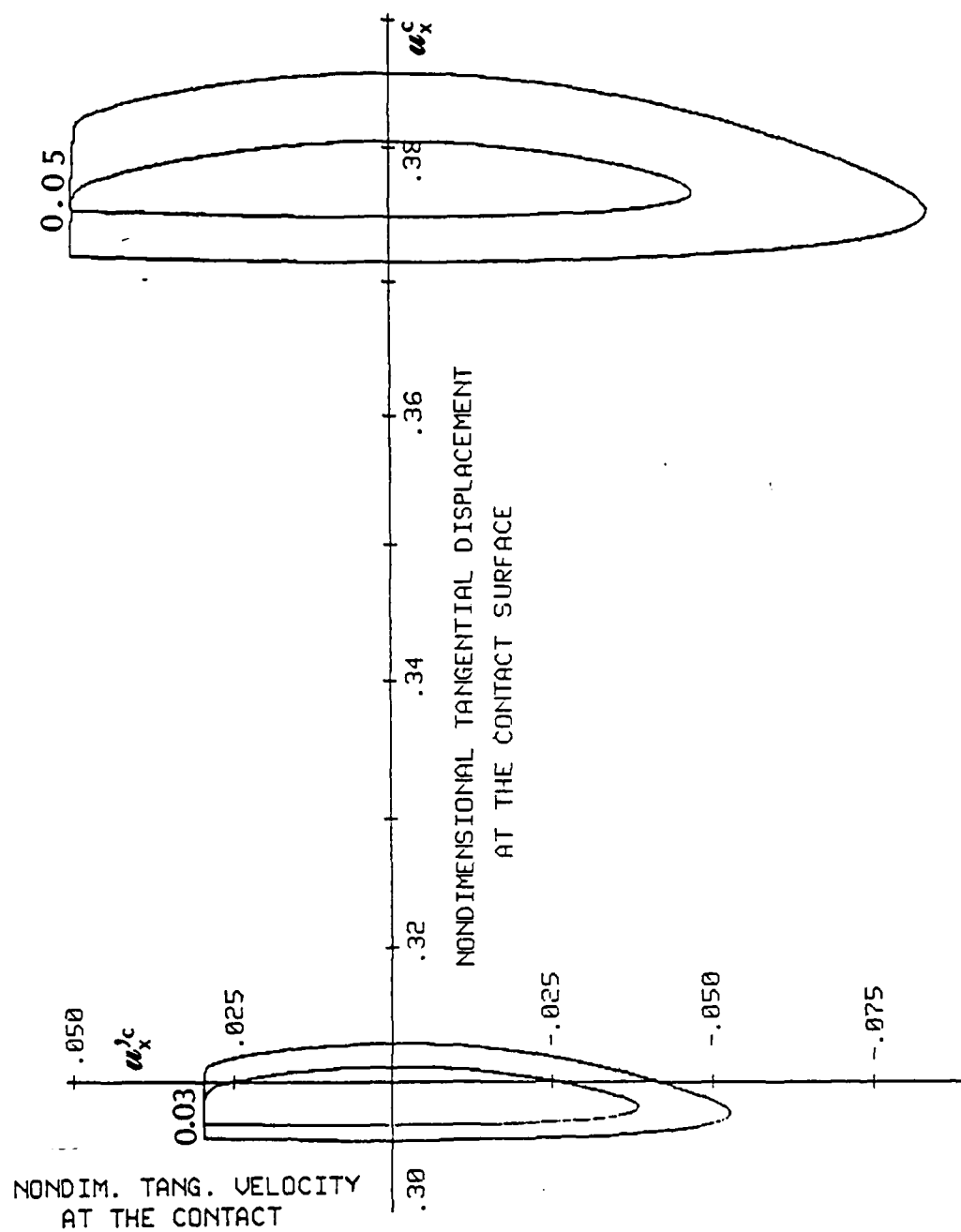


Figure 5.14: Phase plane plots for the tangential motion of the points of the block on the contact surface ($s=0.01$, $z_x=0.001$ and $U_x^C=0.03$ and 0.05).

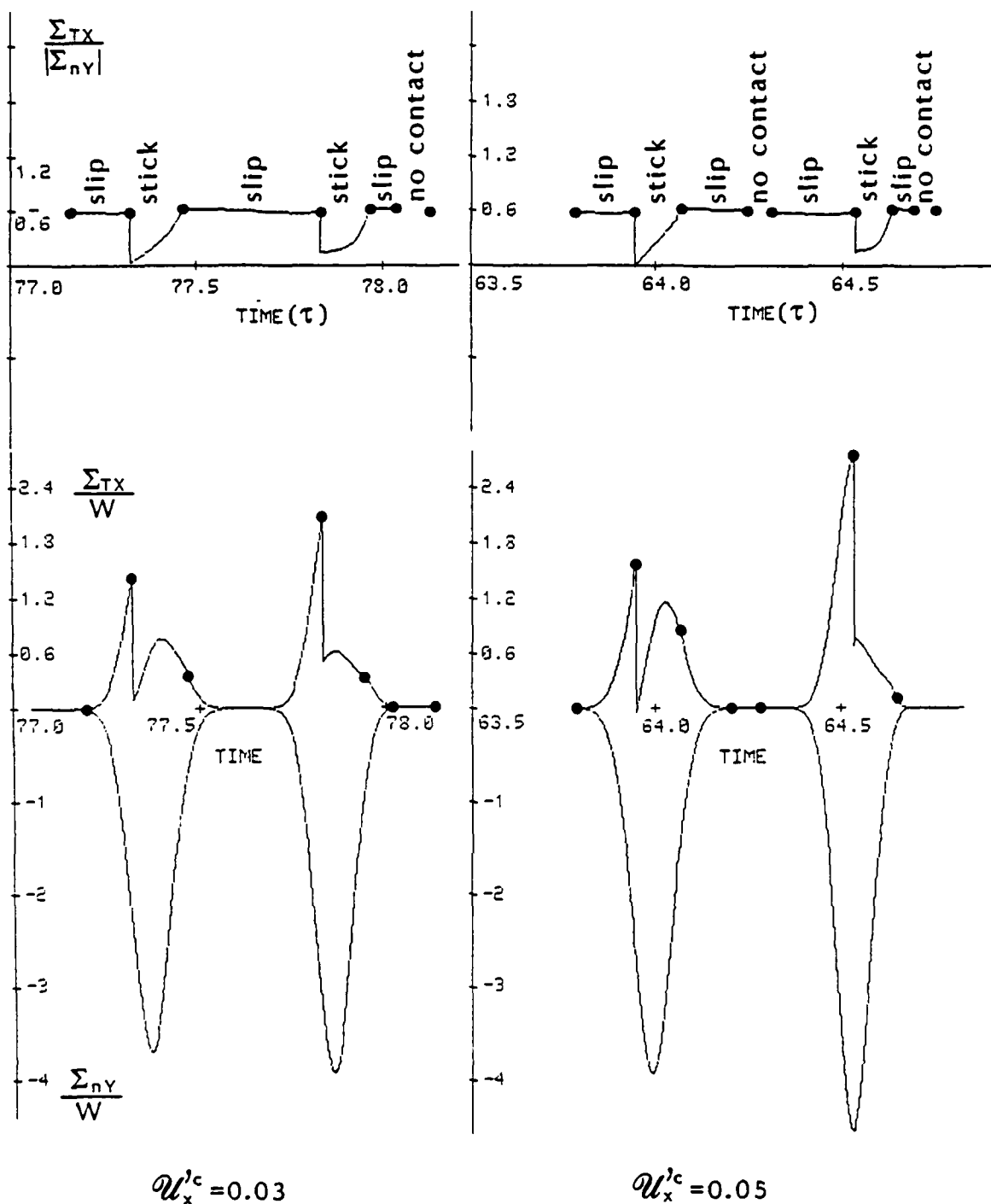


Figure 5.15: Evolution of the normal contact force (Σ_{ny}), the friction force (Σ_{Tx}) and the ratio friction force/absolute value of the normal force, for one cycle of oscillation.

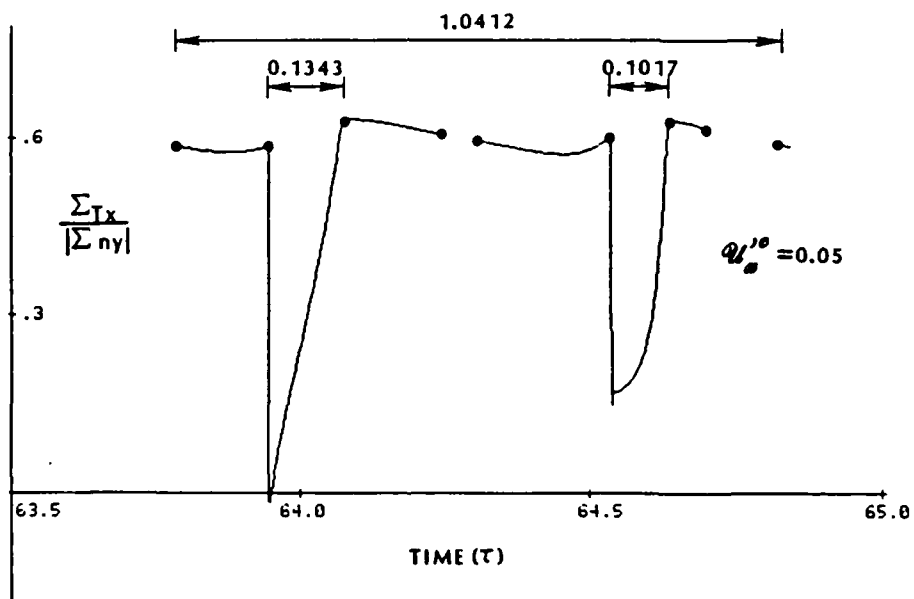
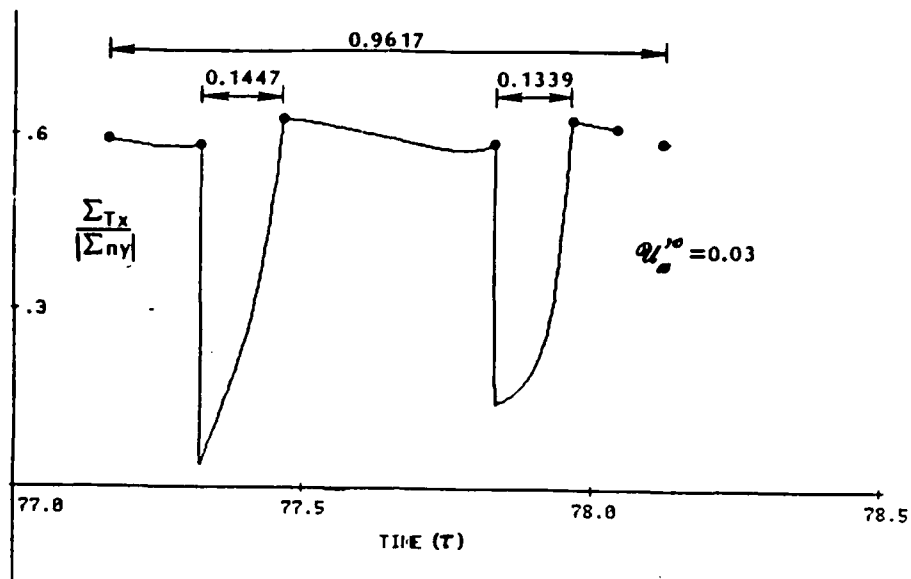


Figure 5.16: The decrease of the time of stick with the increase of the driving velocity ($s=0.01$, $z_x=0.001$ and $U'_C=0.03$ or 0.05).

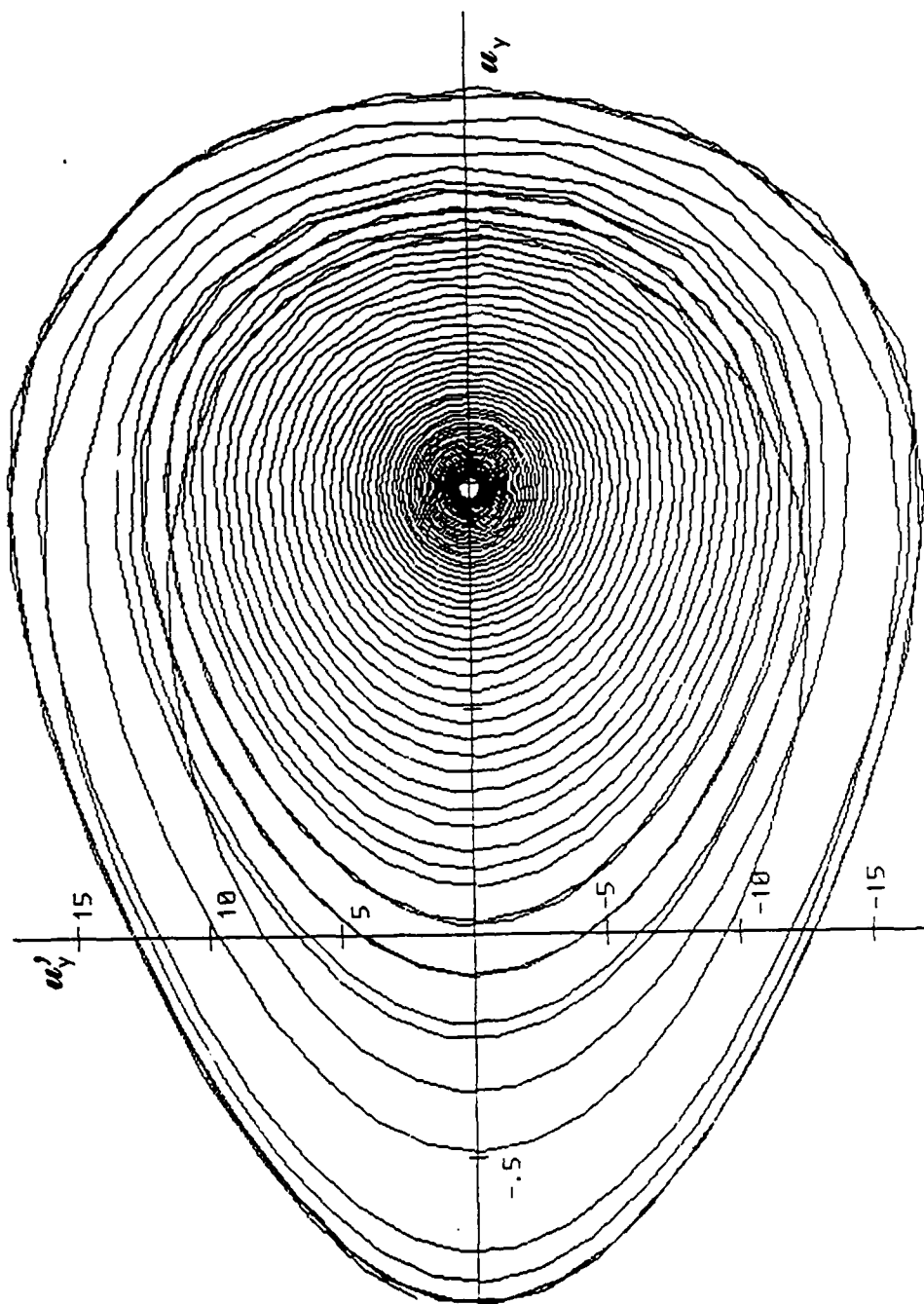


Figure 5.17: Phase plane plot of the decaying portion of the normal oscillation of the center of mass in the course of low-frequency stick-slip motion.

Thus, monitoring the spring elongations, as is often done in friction experiments, case (a) would be perceived as an apparently smooth sliding with a coefficient of kinetic friction smaller than the coefficient of static friction and case (b) would be perceived as a (low-frequency) stick-slip motion.

The role of the normal interface damping. It is clear from the results presented above that for “small” tangential stiffness and damping, the size of the driving velocity plays an important role on the occurrence of low frequency stick-slip motion or apparently smooth sliding motions. It is also clear that when a low frequency stick-slip motion occurs the normal oscillation that accompanies the sliding phase of the stick-slip cycles is damped out when the spring elongation is close to its minimum. The role played by the normal interface damping on the qualitative behavior of the rigid block is made clear in Figs. 5.18 and 5.19. The values used in the computations are those indicated in (5.39) to (5.41) and (5.43) to (5.46), but now we fix $U_x^C = 0.01$ and we vary \hat{z} as indicated on the figures. In Fig. 5.18 it can be observed that, for very small or null interface normal damping ($\hat{z} = 0.002$ or $\hat{z} = 0.000$) an apparently smooth sliding motion is obtained. As the interface normal damping is increased ($\hat{z} = 0.005, 0.01, 0.02, 0.05$) a low frequency stick-slip oscillation is observed. The amplitude of that oscillation decreases with the increase of \hat{z} as shown in Fig. 5.18. The steadiness of the normal oscillation when $\hat{z} = 0.0$ and the intermittency of the normal oscillation when low frequency stick-slip motions occur ($\hat{z} = 0.01$ and $\hat{z} = 0.05$) is made clear in Fig. 5.19. The decrease of the amplitude of the normal oscillation when \hat{z} is increased can also be observed on the same figure. The phase plane plot of the tangential oscillation of the points of the block on Γ_C when $\hat{z} = 0$ is shown in Fig. 5.20.

From the observations above we conclude that some normal interface dissipation is needed for the occurrence of low frequency stick-slip motion. In physical terms, this means that *some plastic deformation (penetration) of the interface must occur at the end of the sliding portion/beginning of the stick portion of the stick-slip cycles and it is that plastic deformation that is responsible for the damping of the normal oscillation.*

Fig. 5.19 suggests another comment: the form of the decay of the normal oscillation at the end of the sliding portion of the stick slip cycles reveals in a clear manner the *viscous* nature of the nonlinear term in (5.1) that is responsible for that dissipation. Experimental observations of the same phenomenon with actual metallic surfaces may suggest more appropriate forms for that normal dissipative term.

Remark 5.3.1. The behaviour reported above for very small normal interface damping is not always observed. When \hat{z} is very small (or null) and with some values of the other

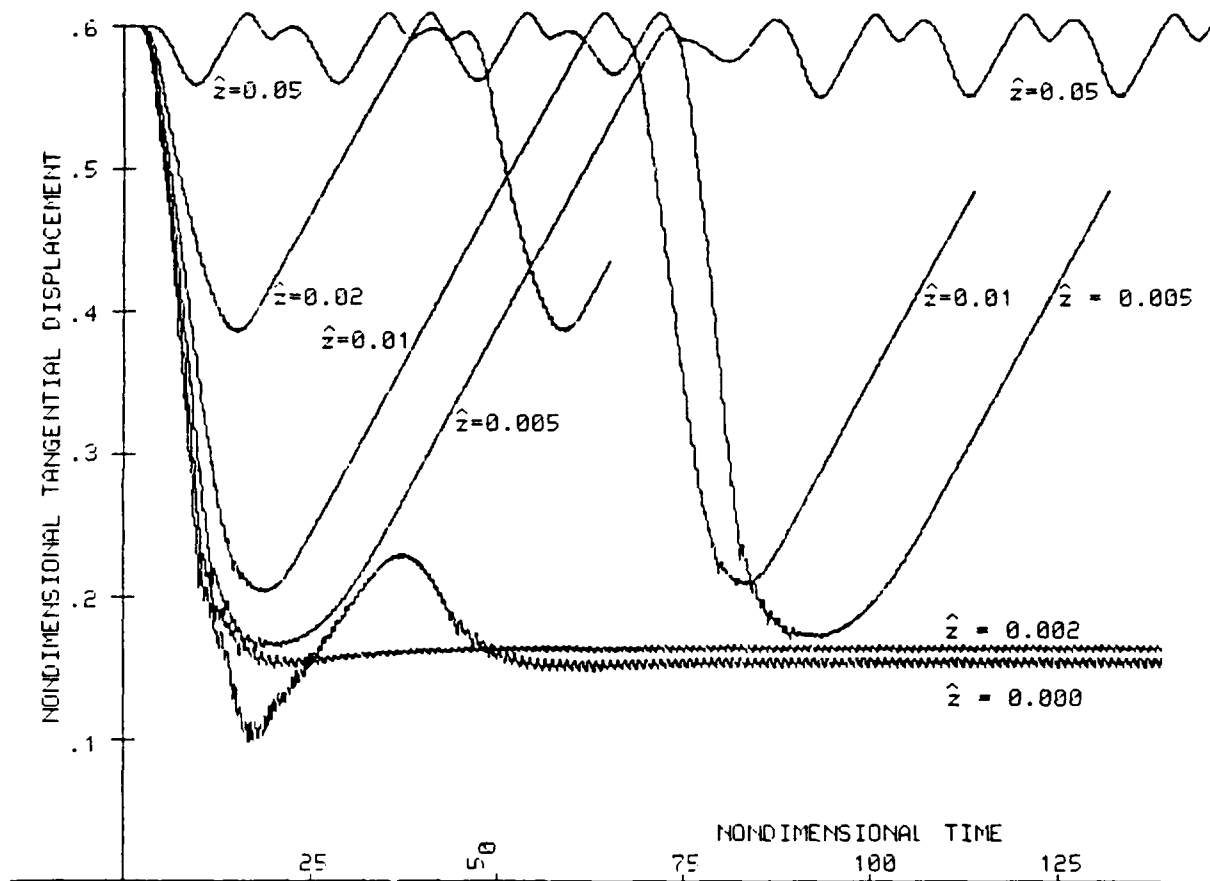


Figure 5.18: Effect of the normal interface damping on the tangential trace.

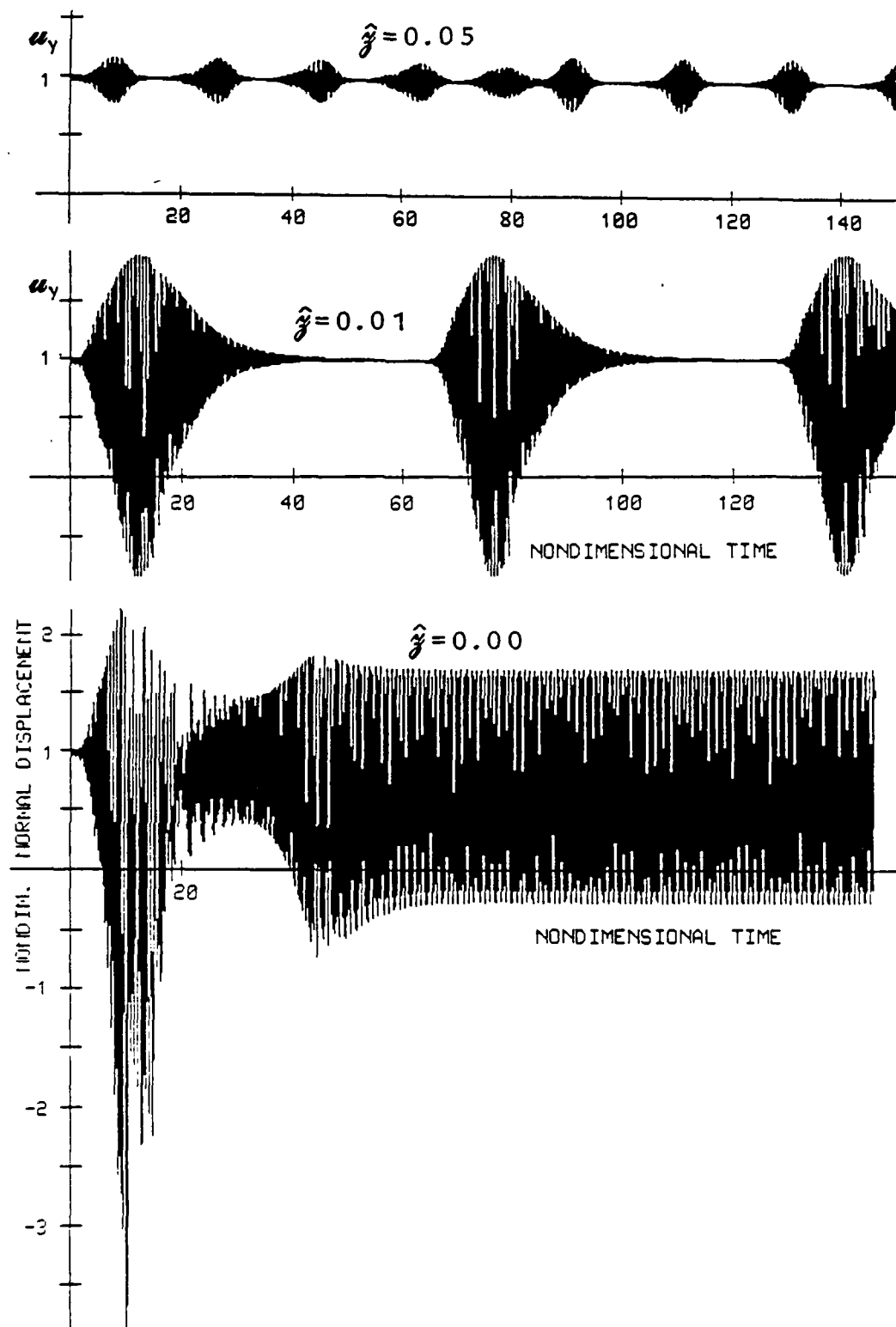


Figure 5.19: Effect of the normal interface damping on the normal oscillations

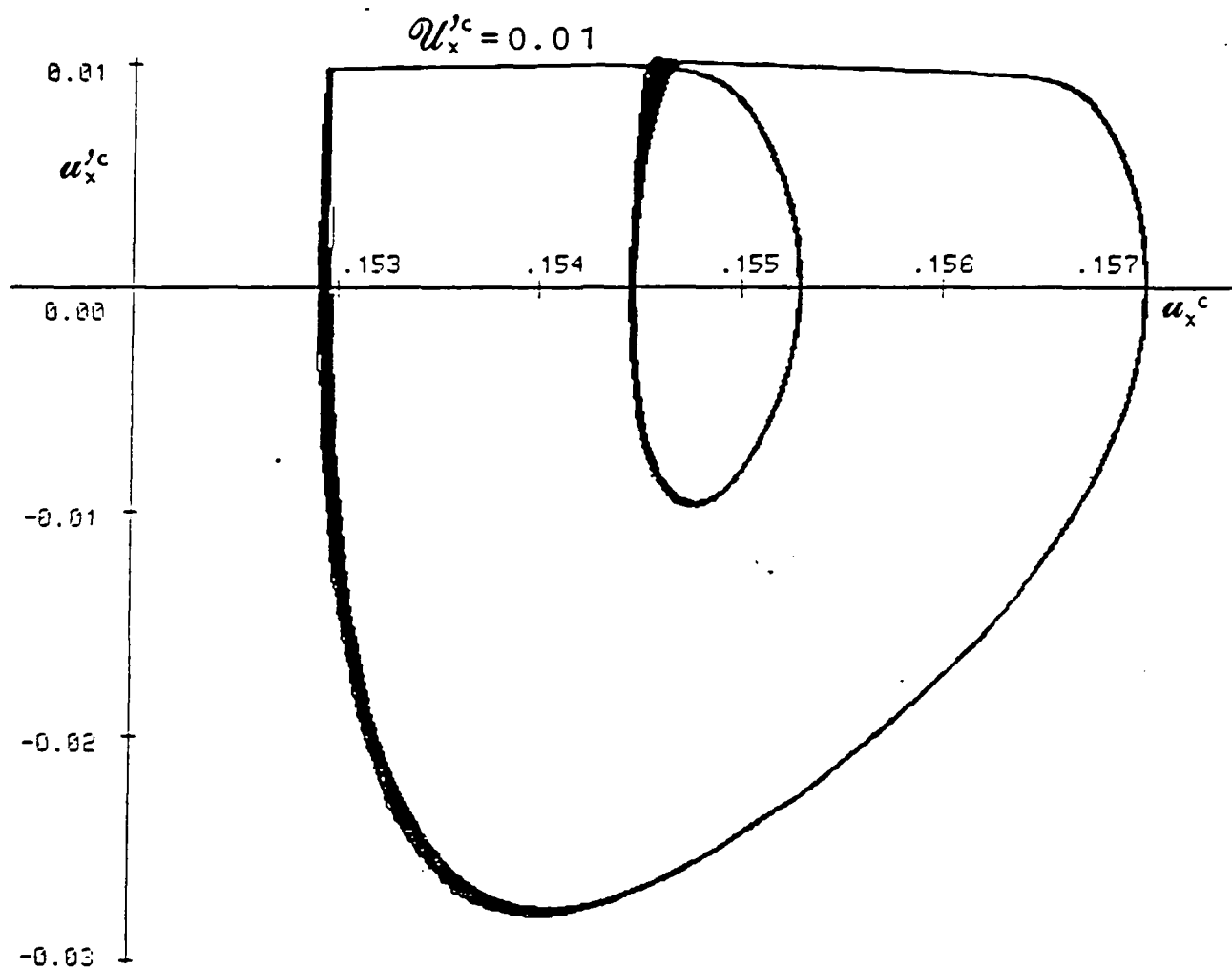


Figure 5.20: Phase plane plot of the steady tangential oscillation of the points of the block on the contact surface for $\hat{z}=0.0$. (Note: not all the computed points are plotted.)

governing parameters the tangential motion may be extremely “irregular” with regions of sliding and sticking alternating in an irregular and unpredictable manner. These irregular frictional traces are qualitatively very similar to some of those reported in [11, 12, 60] and described on the second part of the quotation from [11], in Section 2.1. For further details on these “irregular” numerical traces see [88].

The coefficient of kinetic friction during the slip phase of the low frequency stick-slip motions. Our interest here is to obtain the loops described by the apparent coefficient of kinetic friction when plotted against the sliding velocity during the slip phase of the low frequency stick-slip motions described earlier.

The instantaneous apparent coefficient of kinetic friction has the same definition as in Section 4: it is the instantaneous ratio between the friction force (Σ_{Tx}) and the nominal normal load on the interface (the weight W of the slider). If the tangential linear damping term in (5.31) is neglected, the apparent coefficient of kinetic friction at time τ is given by:

$$\Sigma_{Tx}(\tau)/W = u_x''(\tau) + u_x(\tau)$$

The high frequency oscillations occurring during the slip phase of the low-frequency stick-slip motions are, of course, present in the files with the computed values of $\Sigma_{Tx}(\tau)/W$ and of the sliding velocity ($u_x'(\tau) - U_x^C$). On the other hand, we recall from Section 2.6 that only detailed and specifically oriented experimental studies allowed some authors to detect those high frequency oscillations during the slip phase of stick-slip motions. That was not the scope of most published experimental works and, in some cases, the techniques used smoothed out any high frequency perturbations.

In order to get plots that might be compared with most published experimental results we filtered out the high frequency oscillations in the files of the apparent coefficient of kinetic friction and of the sliding velocity. This was achieved with several passes of those files through a digital filter based on a simple moving average algorithm.

A typical (filtered) plot of the variation of the apparent coefficient of kinetic friction with the sliding velocity during one cycle of the low frequency stick-slip motion is shown in Fig. 5.21. This plot was obtained with the data (5.39), (5.40), (5.41), (5.42) and $U_x^C = 0.01$.

The loop presented in Fig. 5.21 has the clockwise orientation most commonly reported in the literature.

The effect of the stiffness parameter and the driving velocity on the amplitude and frequency of the low frequency stick-slip motion. In order to study the influence

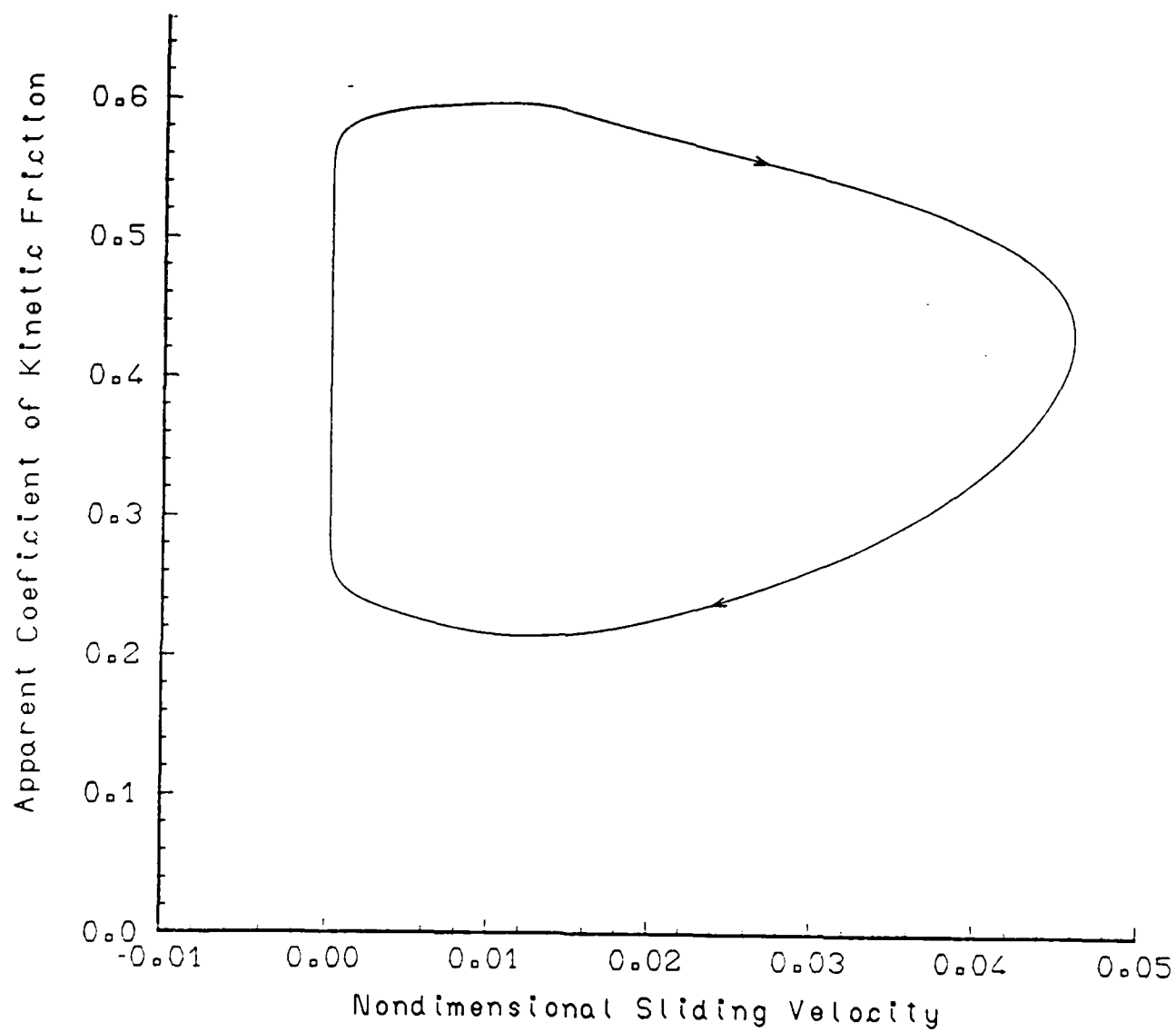


Figure 5.21: Apparent coefficient of friction versus sliding velocity for one cycle of the low frequency stick-slip motion.

of these parameters, we fixed the data (5.39), (5.41), (5.42) and (5.45) and we assigned to s and U_x^C various values in the small ranges for which low frequency stick-slip motion is observed. In Fig. 5.22 and 5.23 we plot the amplitudes of the low frequency stick-slip oscillations against the driving velocities for each of the values considered for the stiffness parameter. The nondimensional amplitude $\Delta \tilde{u}_x$ and the nondimensional driving velocity U_x^C are used in the plot of Fig. 5.22. In Fig. 5.23 different nondimensional variables are used: the amplitude $\Delta \tilde{u}_x/s (= \Delta u_x/Y)$ and the driving velocity $U_x^C/\sqrt{s} (= \dot{U}_x^C/\sqrt{gY})$ where Y is the quantity defined in (5.18) and $g (= W/M)$ denotes here the gravity acceleration. The plots of Fig. 5.23 represent thus the behavior of the dimensional amplitude Δu_x when the changes in s and U_x^C result from changes in K_x and \dot{U}_x^C , respectively, while using the same body and the same contact surface (the same W , the same M and the same Y). *The decrease of the dimensional stick-slip amplitude with the increase of the tangential stiffness can be observed in Fig. 5.23.*

For the range $0.0004 \leq s \leq 0.01$ of the stiffness parameter, *the amplitude of the low frequency stick-slip motion decreases only slightly with the increase of the driving velocity*, which appears consistent with the small slopes of the amplitude-driving velocity plots at small driving velocities, obtained experimentally by several authors ([15, 47, 60]). Also in agreement with some experimental observations [47, 60] *the transition from low frequency stick-slip motion to apparently smooth sliding is abrupt*. For small tangential stiffnesses, the critical driving velocity at which that transition occurs is not very dependent on the stiffness: a (small) decrease of the dimensional critical speed \dot{U}_x^C with the increase of the tangential stiffness is observed in Fig. 5.23, and these facts appear consistent with the experimental observations in [5, 60].

Finally, we observe that the larger stiffness parameters considered ($s=0.025$ and $s=0.035$) correspond to a transition between the range of values of s for which low frequency stick-slip oscillations occur at low driving velocities and the range of values of s for which no low frequency stick-slip motions occur at any driving velocity. For these values of s an increase of the amplitude of the stick-slip motion with the increase of the sliding velocity can be observed in Fig. 5.22, and this is followed also by an abrupt transition to apparently smooth sliding motion at some critical driving velocity. Although increases of amplitude of the stick-slip motion with the driving velocity are not frequently reported in the literature, we observe that in some cases such phenomenon has indeed been observed: as an example we mention the work of Brockley and Ko [17], (in particular see Fig. 13, page 555 of their work, for the small velocity range at which the friction induced oscillation has a saw tooth

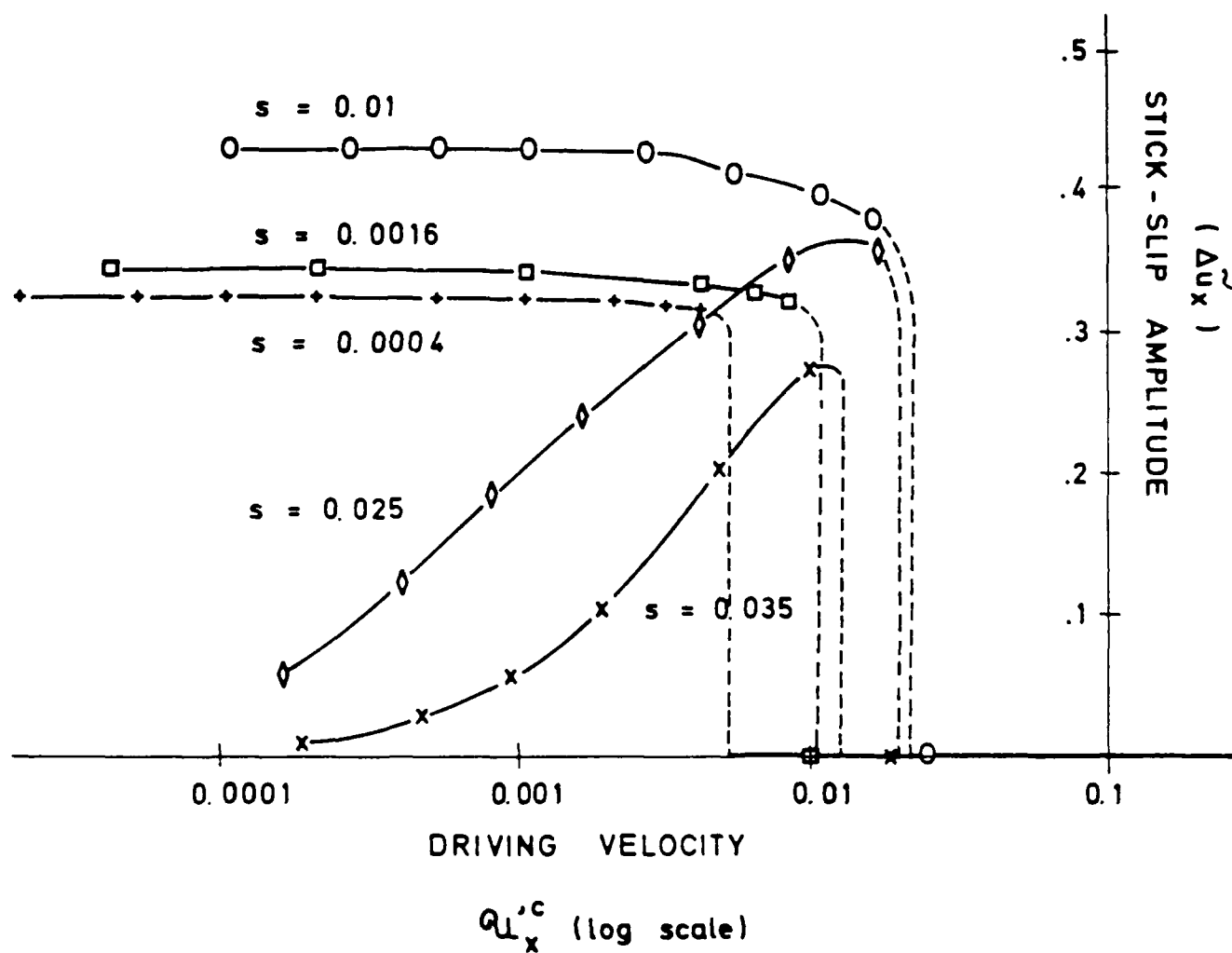


Figure 5.22: Amplitude of the low frequency stick-slip motion $(\Delta \dot{u}_x)$ vs. driving velocity $(U_x^{C'})$.

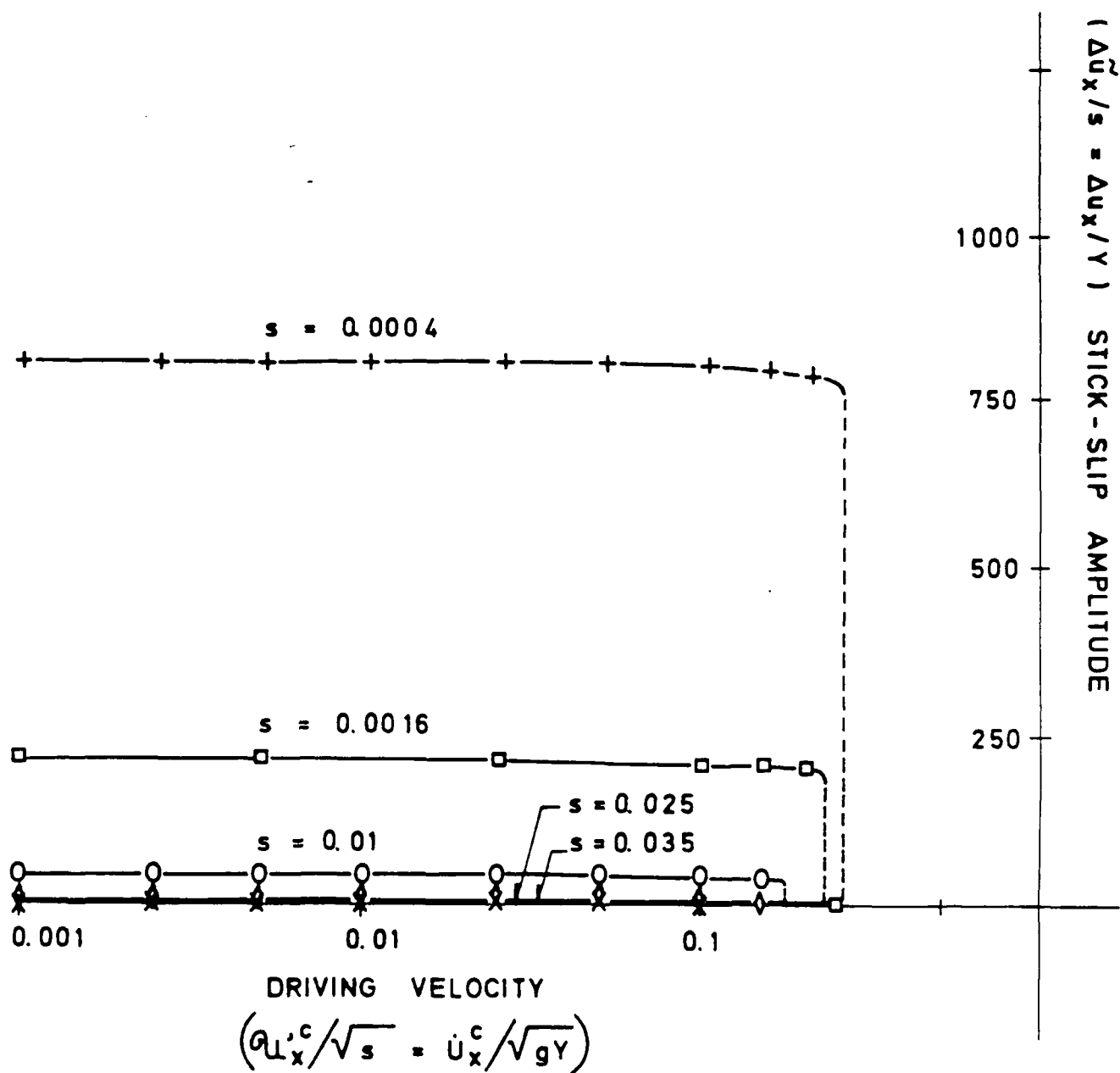


Figure 5.23: Amplitude of the low frequency stick-slip motion vs. driving velocity

wave form).

In Fig. 5.24 a plot of the variation of the nondimensional stick-slip frequency with the nondimensional driving velocity U_x^C is presented. The nondimensional stick-slip frequency is equal to the ratio ω^{ss}/ω where ω^{ss} denotes the dimensional (rad s⁻¹) stick-slip frequency and ω denotes the frequency of the tangential free oscillation (5.30). For the small stiffness parameters ($s=0.0004$; $s=0.0016$ and $s=0.01$) the log-log plot of *the stick-slip frequency versus the driving velocity is essentially linear*, which results from the small dependence of the stick-slip amplitude on the driving speed (Figs. 5.22 and 5.25) and which also bears some similarity with the results of Dokos [28]. In Fig. 5.24 it can be observed that the low frequency stick-slip motion ceases when its frequency is still well below the natural frequency of the tangential motion. Situations of this type have been reported by Kato and Matsubayashi [47] but, as observed in Section 2, it is frequently observed that the low frequency stick-slip motion persists up to the natural frequency of the free tangential motion. At this point it is unclear if, for some values of the governing parameters, our interface model and the rigid block considered here may simulate such a behavior. See Section 5.4 for further related discussions.

The apparent coefficient of kinetic friction-sliding velocity plots for the apparently smooth sliding motions. For “small” stiffness and tangential damping parameters and with a small driving velocity a low frequency stick-slip motion is obtained while, with a large driving velocity, an apparently smooth sliding results. It is also known that two means for obtaining experimentally a smooth sliding at low sliding speeds are: the use of a very stiff tangential spring or a very strong tangential damping. Here we show that our model predicts these behaviors.

In Fig. 5.25 the tangential displacements obtained with the data (5.39, 5.41, and 5.42), a “large” stiffness parameter ($s=0.1$) and various driving velocities are presented. It can be seen that these traces consist of a self-excited oscillation (*without the typical saw-tooth wave form*) about a constant average displacement which corresponds to an apparent coefficient of kinetic friction lower than the coefficient of static friction.

In Fig. 5.26 similar results are shown but now the stiffness parameter has the “small” value $s=0.01$ while the damping parameter has the “large” value $z_r = 10$.

In Figs. 5.27 and 5.28 we plot the apparent coefficients of kinetic friction as a function of the driving velocity (= average sliding speed) for the three cases of apparently smooth sliding that have been considered: small s , small z_r and large U_x^C ; large s , small z_r and arbitrary U_x^C ; small s , large z_r and arbitrary U_x^C . In the second case (large s), since the

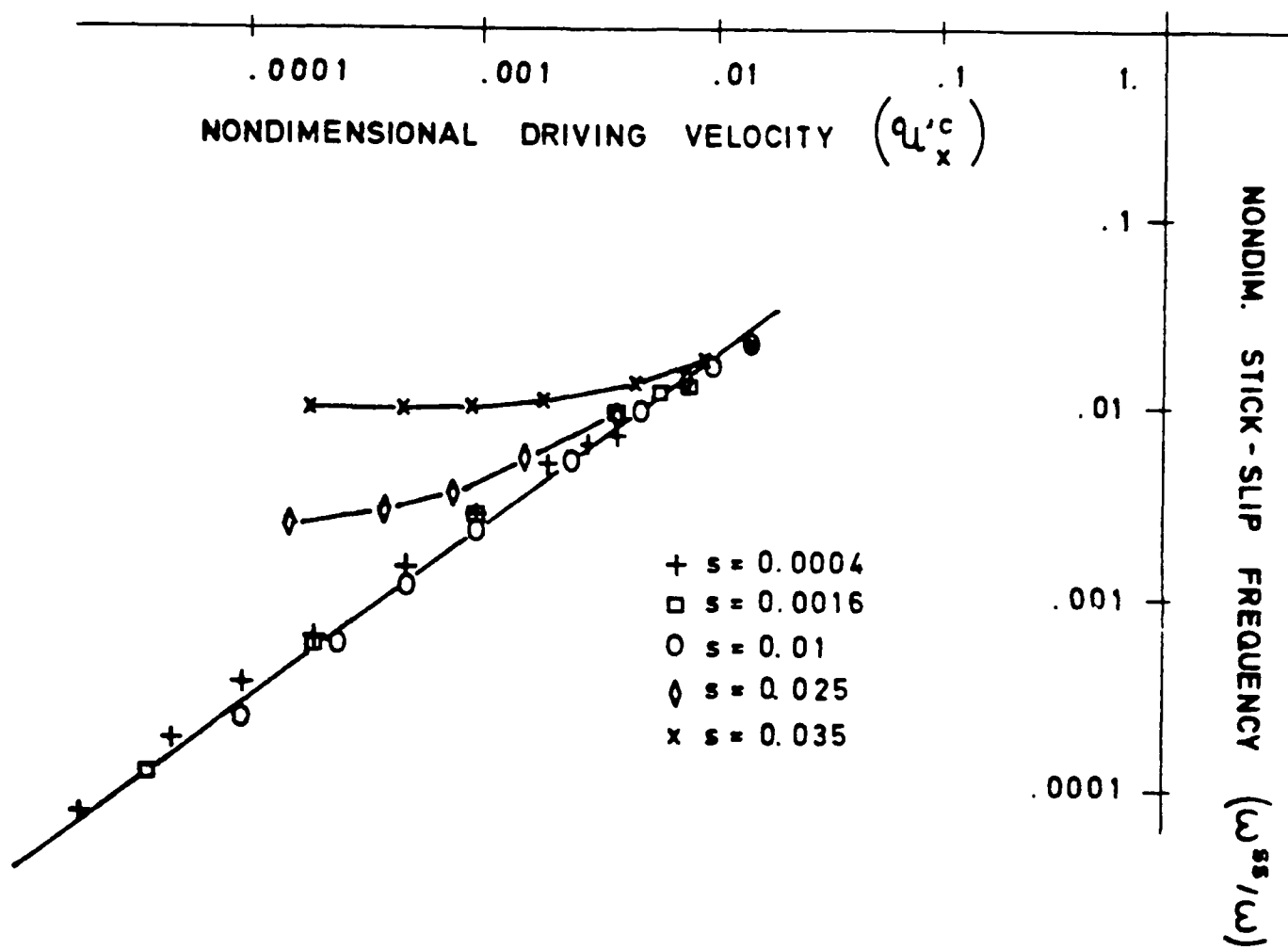


Figure 5.24: Frequency of the low frequency stick-slip motion vs. driving velocity.

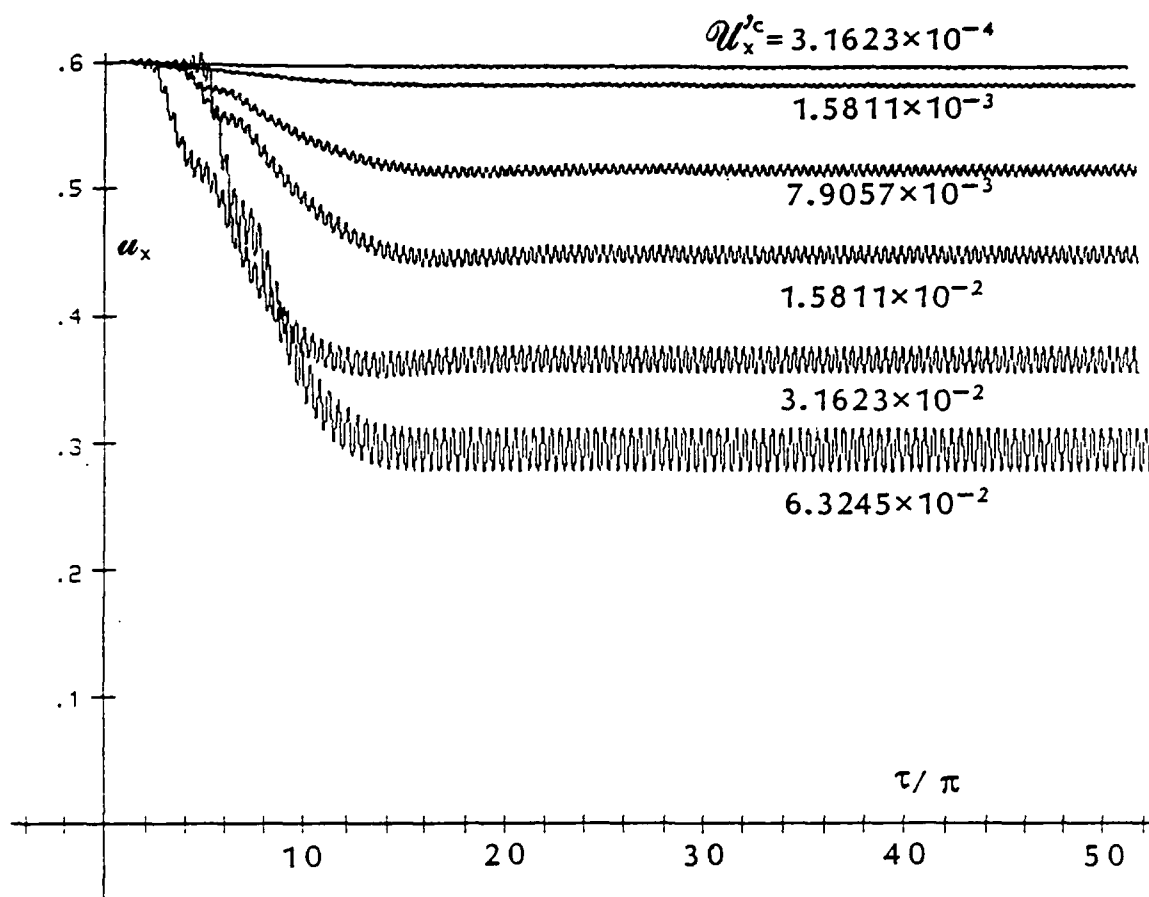


Figure 5.25: Tangential displacement traces for large s , small z_x and various $U_x''^C$ ($s=0.1$, $z_x=0.001$).

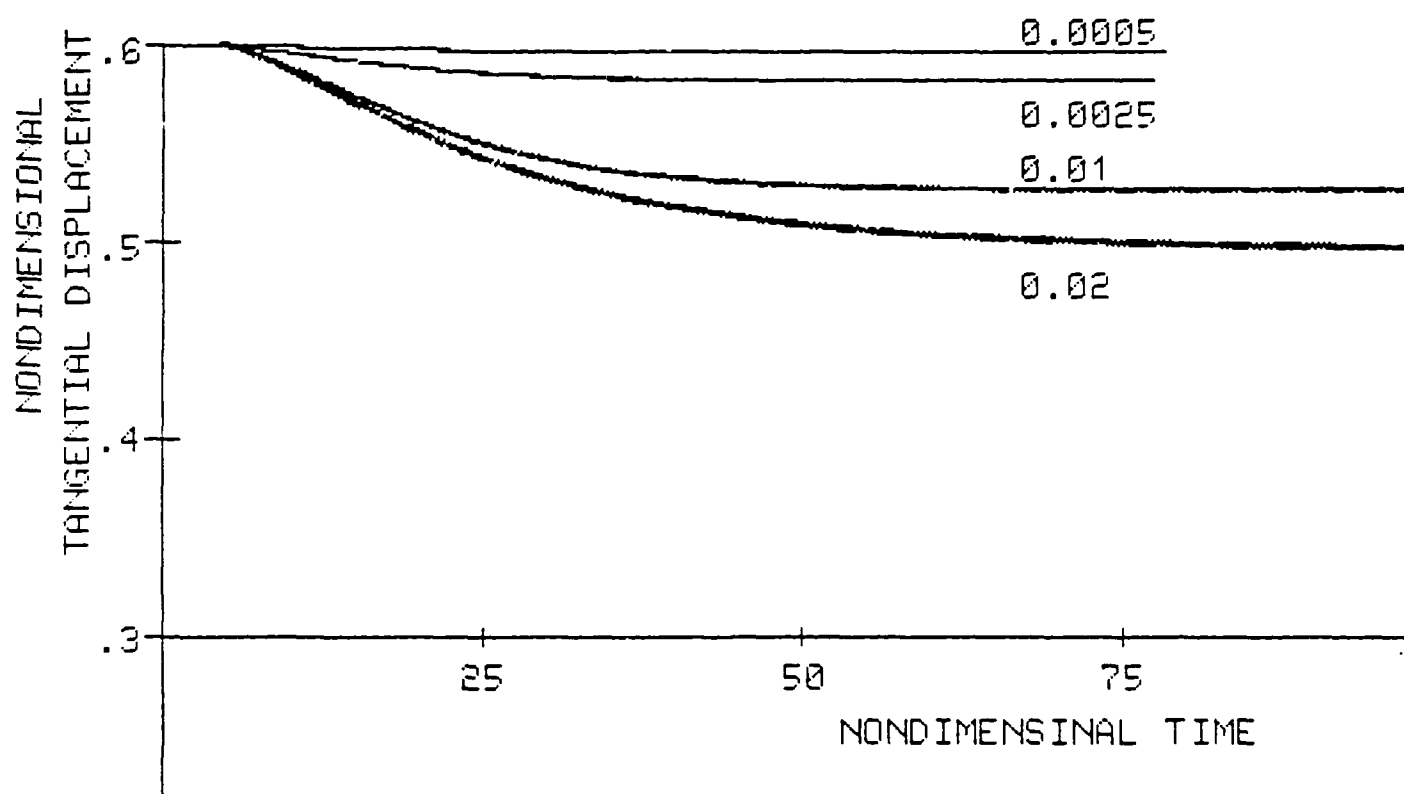


Figure 5.26: Tangential displacement traces for small s , large z_r and various U_r^C ($s=0.1$, $z_r=10$).

nondimensional amplitude of the self-excited oscillations is significant for some speeds, we also indicate those amplitudes on the plot.

The following remarks provide an explanation for apparent coefficients of kinetic friction that are lower than the coefficient of (static) friction and some additional comments:

1. All the low apparent coefficients of kinetic friction in our computations result from the occurrence of a period of stick during each period of oscillation of the body (see Figs. 5.14, 5.15, 5.16, 5.30 and 5.31). The ratio (friction force/normal contact force), when the body sticks, is smaller than the coefficient of friction, so that the time average of the friction force is smaller than the product of the coefficient of (static) friction and the time average of the normal contact force (the weight of the body). *This is precisely the high-frequency stick-slip mechanism proposed by Budanov, Kudinov and Tolstoi in [18] (recall Section 2.6).*
2. If the driving velocity is sufficiently small that periods of stick are possible during each cycle of oscillation the following effects can be observed:
 - (a) The periods of stick occur in the portion of each cycle for which the normal contact force has larger absolute values (see Figs. 5.15 and 5.31)
 - (b) As should be expected, a discontinuity of the friction force occurs at the instant of each transition from slip to stick (see Figs. 5.15, 5.16, and 5.31)
 - (c) For otherwise similar conditions, larger driving velocities imply larger self-excited oscillations (compare in each of Fig. 5.13, 5.14, 5.29 and 5.30 the results obtained with different driving velocities)
 - (d) Also, for otherwise similar conditions, larger driving velocities imply smaller periods of stick, relative to the total period of one oscillation (see Fig. 5.15, 5.16 and 5.31).
3. Increases of apparent coefficient of kinetic friction with the increase of driving velocity result essentially from the effect (d) above (see Fig. 5.16). This is (with the extra complexity inherent to having a three-degree-of-freedom system rather than a two-degree-of-freedom system) what Budanov, Kudinov and Tolstoi [18] suggested to explain apparent coefficients of kinetic friction *increasing* with the average sliding speed (recall mechanism (II) in Section 2.6).

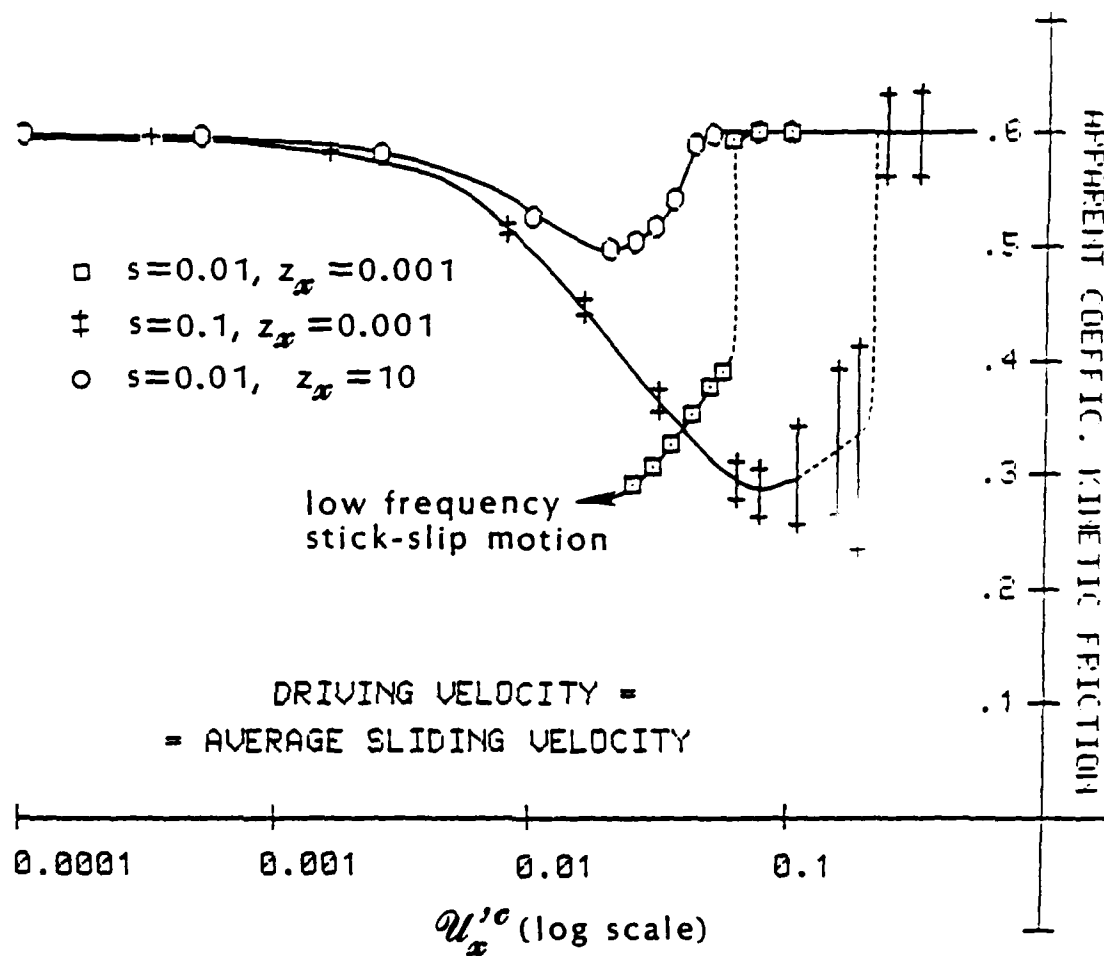


Figure 5.27: Apparent coefficient of kinetic friction vs. average sliding velocity (log scale).

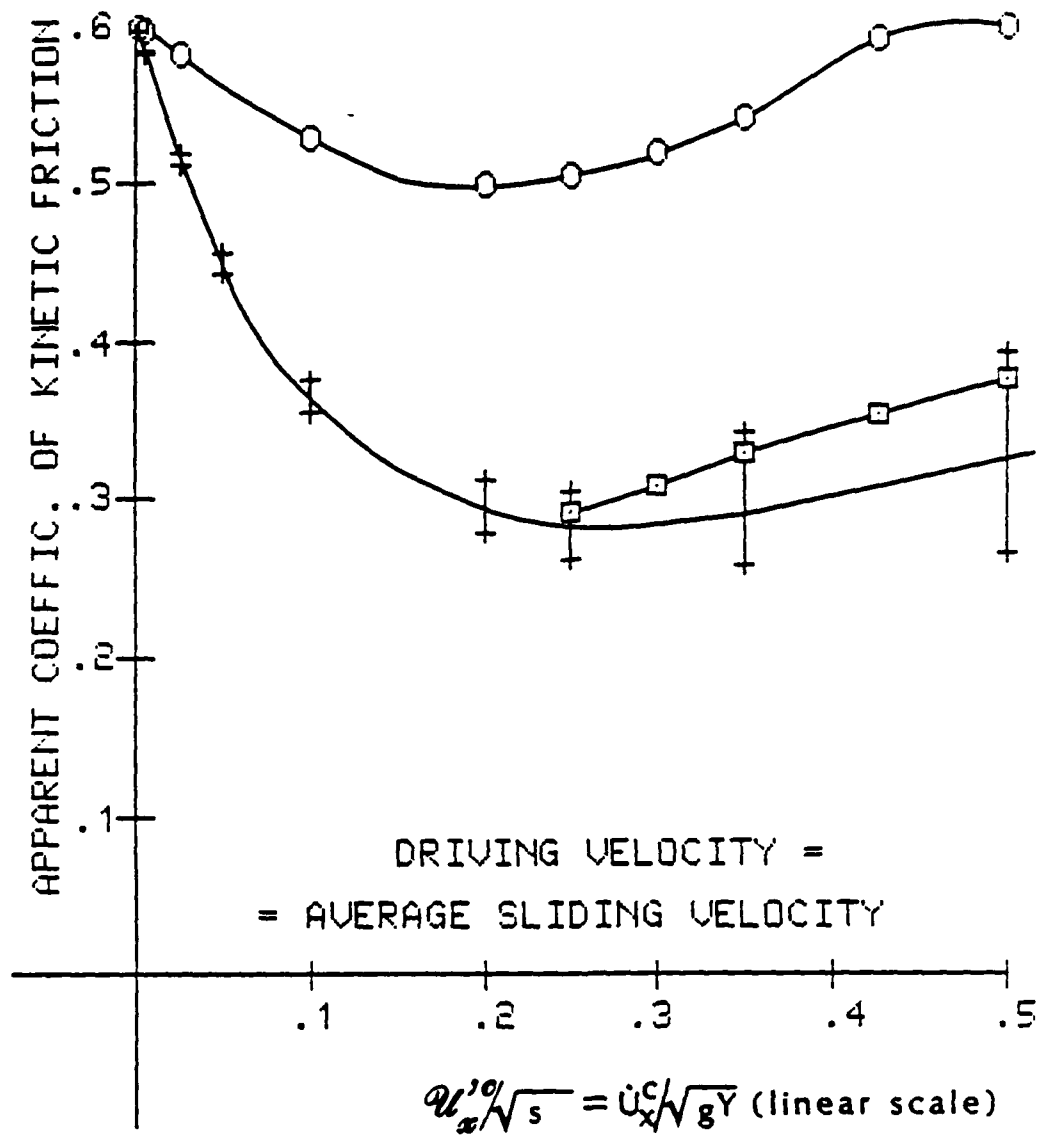


Figure 5.28: Apparent coefficient of kinetic friction vs. average sliding velocity (linear scale).
 ([·]) $s=0.01$, $z_x=0.001$; (+) $s=0.1$, $z_x=0.001$; (○) $s=0.01$, $z_x=10$.

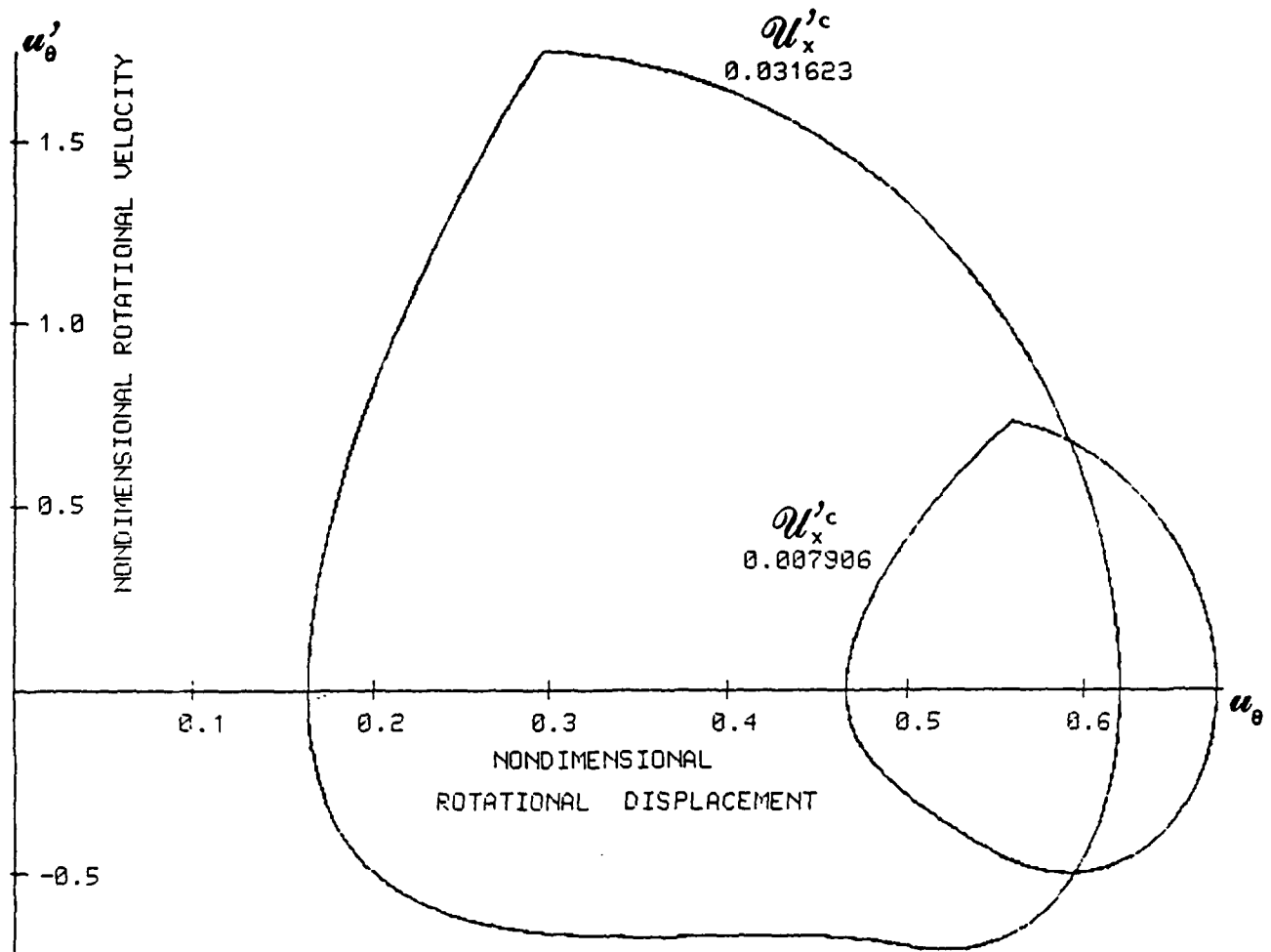


Figure 5.29: Phase plane plots for the rotational motion ($s=0.1$, $z_r=0.001$ and $U_x'^c=0.007906$ or 0.031623). Final self-excited oscillation (the last 2000 steps computed).

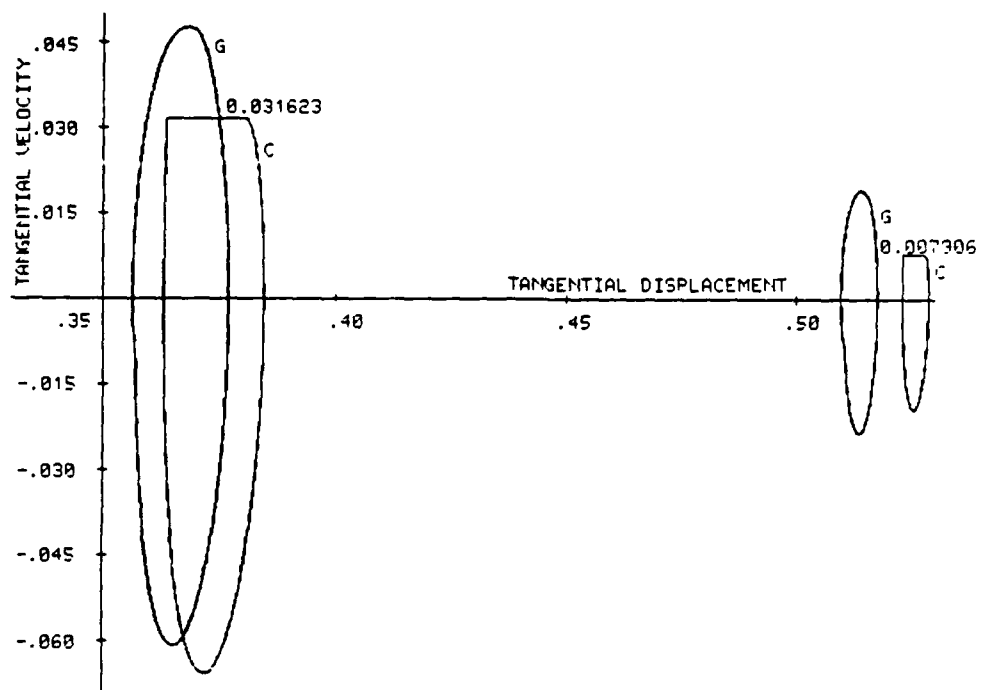


Figure 5.30: Phase plane plots for the tangential motion of the center of mass (G) and the points of the block on the contact surface (C) ($s=0.1$, $z_x=0.001$ and $U_x^C = 0.007906$ or $U_x^C = 0.031623$). Final self-excited oscillation.

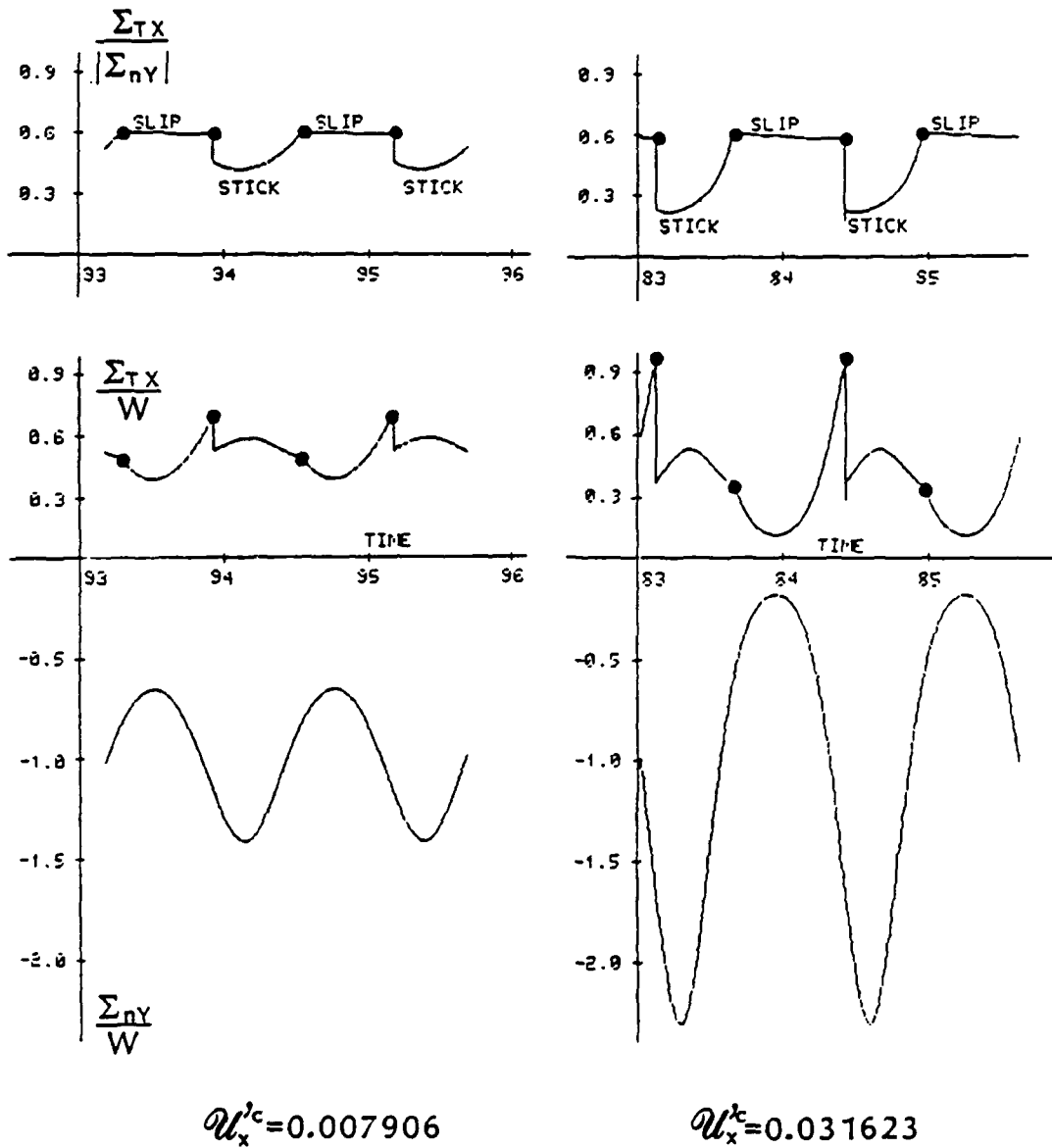


Figure 5.31: Evolution of the normal contact force (Σ_{ny}), the friction force (Σ_{Tx}) and the ratio friction force/absolute value of the normal force, for two cycles of oscillation ($s = 0.1$, $z = 0.001$ and $U_x^c = 0.007906$ or 0.031623).

4. Decreases of apparent coefficient of kinetic friction with the increase of driving velocity are associated with the effects (a) , (b) and (c) above. Compare in Fig. 5.31, for the two driving velocities considered, the minimum values of the friction force (sliding phase), the size of the discontinuities in friction force, and the values of the friction force at the points of maximum normal force (stick phase).
5. For a sufficiently large driving velocity, no stick state occurs during the oscillation so that, since we do not take into account any thermal softening effects, the average coefficient of kinetic friction is equal to the coefficient of static friction.
6. In the case of the large tangential damping ($s=0.01$, $z_x = 10$) we can describe our results using precisely the words of Tolstoi in [74]: "*sufficiently heavy damping of tangential vibrations alone could suppress these vibrations (the low frequency stick-slip motions) but failed to affect the negative slope of the friction-velocity curve.*" In this respect we note that the decoupling (5.24, 5.25) of the characteristic equation for the eigenvalue problem (5.15) implies that, for the present geometry, the introduction of heavy external tangential damping does not affect at all the instability of the steady-sliding. For viscous stabilization of an unstable steady-sliding sufficiently strong normal and rotational damping are required (see Oden and Martins [56]). This, of course, agrees with the observations of Tolstoi summarized in Section 2.6.
7. Fig. 5.30 contains an important warning to experimental researchers of sliding friction: a small, apparently negligible oscillation on the tangential displacement trace at the point where it is being recorded (in our case the center of mass of the block or the tangential spring) may be the subtle manifestation of a stick-slip motion on the contact surface. A very misleading point is the fact that the wave form of the recorded oscillation may be very different from the typical saw-tooth wave form.
8. $\mu - v_T$ plots for apparently smooth sliding motions obtained with different "experimental apparati" (different s and z_x) may be clearly distinct (see Figs. 5.27 and 5.28).
9. The initially decreasing portions of the $\mu - v_T$ curves in Fig. 5.27 and Fig. 5.28 are qualitatively similar to those experimentally obtained by Rabinowicz, Fig. 4.45, page 101 of [60] and by Bell and Burdekin, Fig. 4, page 1078 of [6], or by Watari and Sugimoto, Fig. 1, page 40 of [86], respectively.

5.4 Apparent Reductions of Static Friction Due to Normal Perturbations

5.4.1 Introduction

In Section 5.2 we studied the linear stability of the steady-sliding equilibrium and showed that for some range of the governing parameters, steady-sliding could not be stable: consequences of this on the dynamic behavior of the slider were shown in Section 5.3. Here we are interested in analyzing the effect of normal perturbations on the loading path of the slider — the stick portion of the stick-slip cycles. In a sense, we should like to determine some sort of stability statement concerning that loading path: if a perturbation is introduced while the body sticks, will the body “recuperate” from such a perturbation and keep stuck until the tangential displacement attains the value at which the unperturbed system initiates sliding or, on the contrary, will the body initiate sliding “prematurely” with an apparent coefficient of static friction lower than the true one? We also want to know what effect the driving velocity (or, equivalently, the rate of application of the tangential force) has on the apparent value of the coefficient of static friction at which the perturbed system initiates sliding.

5.4.2 Numerical results

The results reported in this section were obtained with the common data (5.39, 5.40, 5.41, and 5.42).

First we ran our program with the driving velocity successively assuming the values $U_{xk}^C = 2.5 \times 10^{-3}, 5 \times 10^{-3}, 1 \times 10^{-2}, 2.5 \times 10^{-2}, 5 \times 10^{-2}, 1 \times 10^{-1}, 2.5 \times 10^{-1}$, and with the initial conditions $\bar{u}_{x0} = \bar{u}_{y0} = \bar{u}_{x1} = \bar{u}_{y1} = 0, \bar{u}_{x1} = U_{xk}^C, \bar{u}_{y0} = 1$, i.e., the body initiates its motion stuck with the moving surface and no normal perturbation is introduced either at start-up or during the subsequent motion. In the subsequent motion, the body remains stuck until the nondimensional tangential displacement attains a value of the order of f at which sliding initiates. More precisely, sliding always initiates at a value of u_x somewhat in excess of f , due to the inertia acquired by the slider during the stick phase: this excess is almost imperceptible ($\sim 10^{-4}$) for the smallest driving velocities and clearly noticeable ($\sim 10^{-2}$) for the largest velocities considered. For each of the driving velocities considered, the maximum tangential displacement of the slider $u_{xk}^{max} = u_x^{max}(U_{xk}^C)$ is recorded. Then, for each of the driving velocities U_{xk}^C considered, the program is successively run starting at the time τ_{ki} at which the unperturbed tangential displacement was equal to successively decreasing values $u_{xi}=0.575; 0.550; 0.525; \dots$. The initial conditions for these successive runs are the

following: the tangential displacement and velocity, the normal velocity, and the rotational displacement and velocity ($u_{xi}, u'_{xki}, u'_{yki}, u_{\theta ki}, u'_{\theta ki}$, respectively) are precisely the same as those of the unperturbed system at the time τ_{ki} ; the normal displacement is made equal to the unperturbed normal displacement at $\tau_{ki}(u_{yki})$ plus a perturbation ${}^p u_{yj}$ successively equal to -0.05, -0.10, -0.15, -0.20. In other words, for each driving velocity U_{xk}^C we introduce a normal displacement perturbation ${}^p u_{yj}$ at the time τ_{ki} at which the unperturbed tangential displacement was equal to u_{xi} .

The numerical results obtained show that *normal perturbations can produce apparent reductions of the coefficient of static friction and that a fixed level of perturbation has a "destabilizing" effect that increases with the increase of the driving velocity*, i.e., a certain amount ${}^p u_{yj}$ of normal perturbation at a fixed u_{xi} is more likely to produce a "premature" sliding if the driving velocity is large than if it is small. We illustrate this in Figs. 5.32 and 5.33.

In Fig. 5.32 we can observe that a normal perturbation ${}^p u_{yj} = -0.05$ at $u_{xi} = 0.575$ is capable of producing a premature sliding for the larger driving velocities $U_{xk}^C = 5 \times 10^{-3}, 2.5 \times 10^{-3}$, and 1×10^{-3} , but, for $U_{xk}^C = 5 \times 10^{-4}$ the system, after short periods of sliding immediately after the perturbation, sticks again and the maximum tangential displacement is essentially equal to the unperturbed u_{xk}^{\max} . In Fig. 5.33 a normal perturbation ${}^p u_{yj} = -0.1$ at $u_{xi} = 0.55$ produces a "premature" sliding for $U_{xk}^C = 1 \times 10^{-2}, 5 \times 10^{-3}$ and 2.5×10^{-3} but, for $U_{xk}^C = 1 \times 10^{-3}$ the body "recuperates" from the perturbation.

Additional qualitative information can be obtained from the numerical results by searching, for each fixed pair $(U_{xk}^C, {}^p u_{yj})$, the minimum value of u_{xi} at which the normal perturbation considered originates a "premature" sliding. Since actual perturbations are somehow distributed throughout time, it is reasonable to expect that the accumulation of the destabilizing effects of the perturbations will produce a "premature" sliding shortly after the above mentioned minimum of u_{xi} is achieved along the (stick) loading path. A difficulty, however, arises: since the unperturbed system has its maximum tangential displacement at a value u_{xk}^{\max} that is not equal to f and that increases with U_x^C we cannot decide whether "premature" sliding of the perturbed system occurs or not by simply comparing the maximum value of the perturbed tangential displacement with f . In the results reported below we use the following *arbitrary* criterion: "premature" sliding occurs if the maximum tangential displacement of the perturbed system (${}^p u_{xk}^{\max}$) is at least one percent smaller than the corresponding maximum tangential displacement of the unperturbed system (u_{xk}^{\max}) for the same driving velocity U_{xk}^C , i.e., if $(u_{xk}^{\max} - {}^p u_{xk}^{\max})/u_{xk}^{\max} \geq 0.01$. We also note that, since we

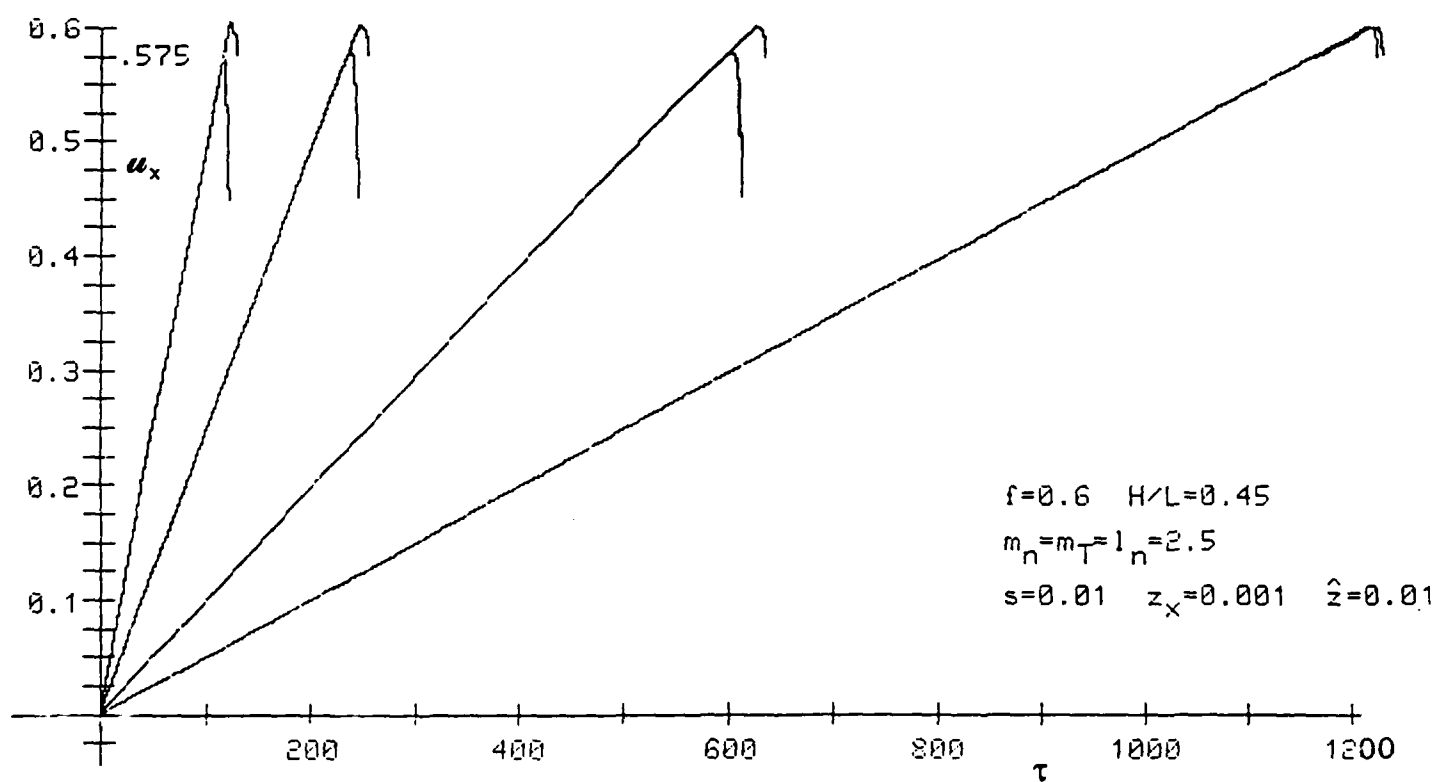


Figure 5.32: Unperturbed and perturbed tangential displacements for a normal perturbation $^p u_{yj} = -0.05$ at $u_{xi} = 0.575$. Driving velocities $U''^C_{xk} = 5 \times 10^{-3}$, 2.5×10^{-3} , 1×10^{-3} and 5×10^{-4}

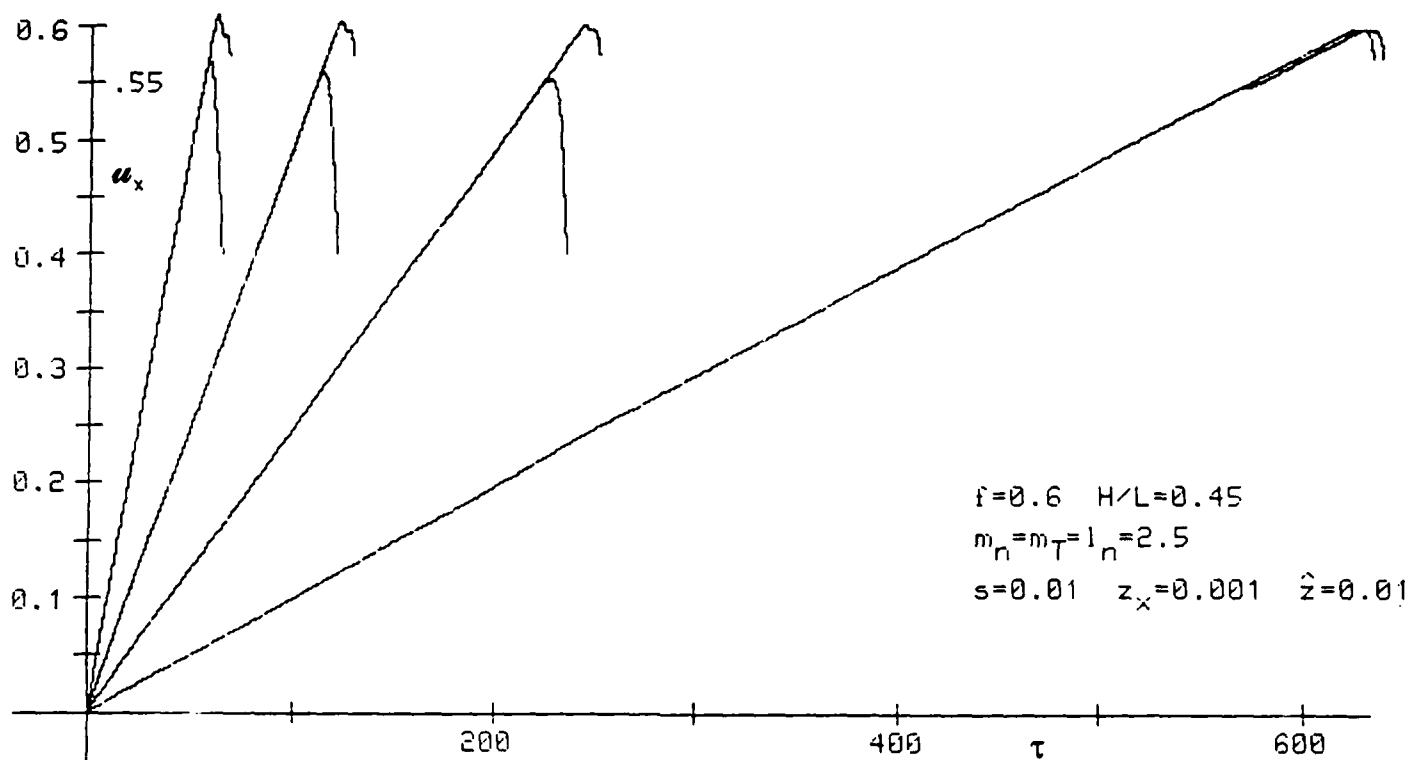


Figure 5.33: Unperturbed and perturbed tangential displacements for a normal perturbation $p_{u_{yj}} = -0.10$ at $u_{xi} = 0.55$. Driving velocities $U'_{xk} = 1 \times 10^{-2}$, 5×10^{-3} , 2.5×10^{-3} and 1×10^{-3}

only considered perturbations at discrete locations $u_{xi} = 0.575, 0.550, 0.525, \dots$ to say, for instance, that $u_{xi} = 0.550$ is the minimum value of u_x at which some perturbation produces "premature" sliding only means that such perturbation at $u_{xi} = 0.575$ and 0.550 produces "premature" sliding while the same perturbation at $u_{xi} = 0.525, 0.500, \dots$ does not produce "premature" sliding.

With these conventions in mind we summarize the results obtained in Fig. 5.34. In addition to observations made earlier we observe, as might be expected, that *large perturbations produce a "premature" sliding more easily than small ones*. We observe also that *sufficiently large normal perturbations can produce apparent reductions of the coefficient of static friction even for the smallest driving velocities considered*, and this is consistent with the experimental observation by Toistoi [74] that, as soon as an external normal damping is imposed on a slider, the force of static friction increases "no matter how slowly the pulling force has been increasing."

Using the same driving velocities U_{xk}^C , the same locations u_{xi} for the perturbations and the same conventions, an additional study was done to analyze the effect of the stiffness of the restraining horizontal arm on the apparent reductions of static friction. In these computations we used the data (5.39), (5.41), (5.42), a fixed level of perturbation $p u_y = -0.15$ and the values 0.0016, 0.01 and 0.035 for the stiffness parameter s . For each value of s the driving velocities U_{xk}^C were sufficiently small that the subsequent motions were low frequency stick-slip oscillations.

For each value of s we plot in Fig. 5.35 the minimum nondimensional tangential displacements (u_{xi}) along the stick loading path at which the normal perturbation $p u_y = -0.15$ produces a "premature" sliding *versus* a measure of the time of stationary contact (τ_{ki}/\sqrt{s}). This measure allows us a qualitative comparison of the computed results with some experimental ones which were obtained by varying the tangential stiffness K_x and the driving velocity U_x^C while keeping the same slider and the same interface ($g = W/M$ and $Y = (W/c_n BL)^{1/m_n}$ are constants). In these circumstances the nondimensional quantity $\tau_{ki}/\sqrt{s} = t_{ki}\sqrt{g/Y}$ is proportional to the physical time (t_{ki}) spent on the stick loading process up to the point where the perturbation originates a "premature" sliding.

In Fig. 5.35 it can be observed that *for each time of stationary contact, the apparent coefficient of static friction increases with the increase of the tangential stiffness*. Such conclusion is in qualitative agreement with the experimental observations of Bell and Burdekin (Fig. 6a, page 552 of [5]), exception made for the fact that (unlike the conclusions of the survey [64]) those authors do not report a constant coefficient of static friction for large

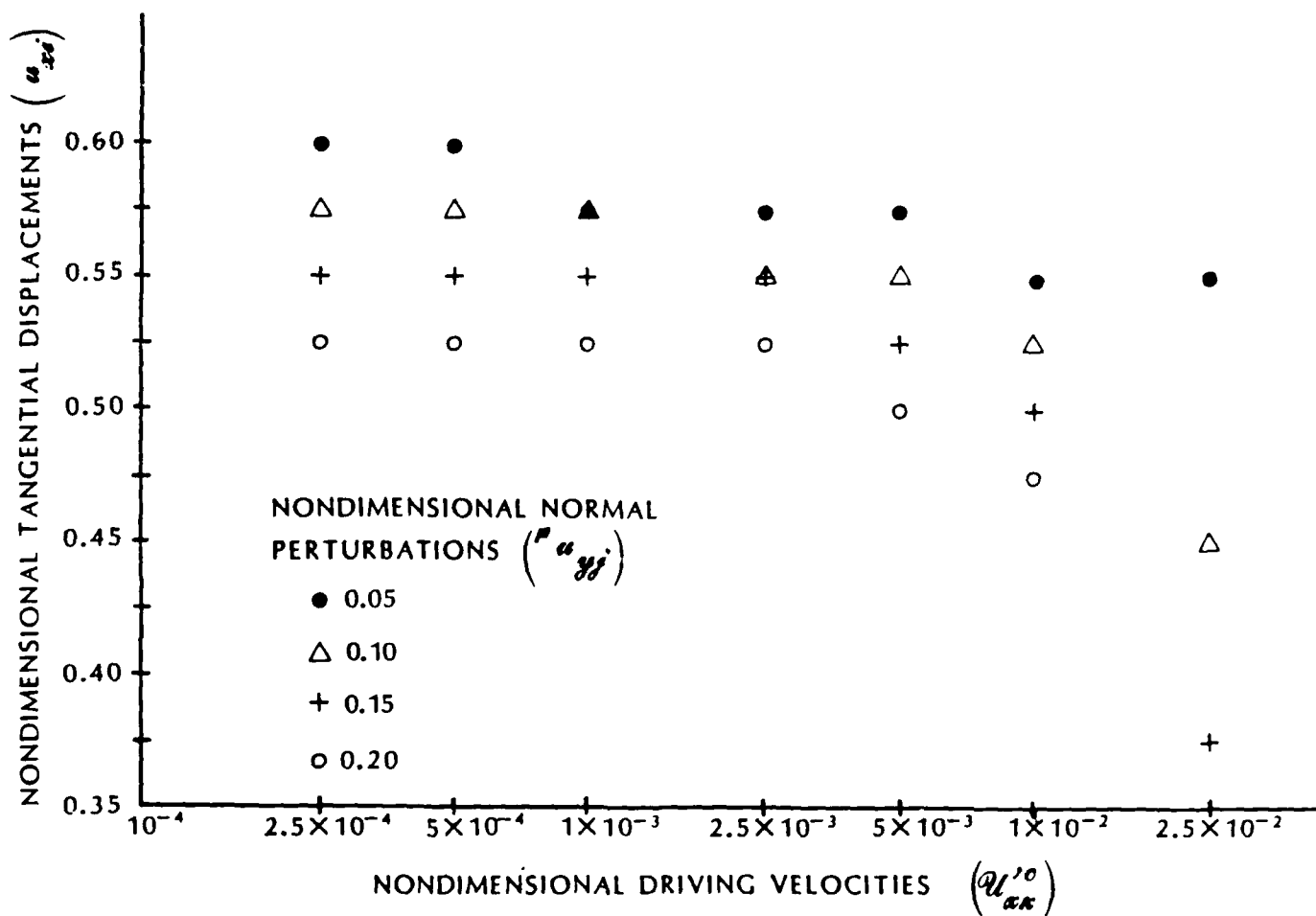


Figure 5.34: Minimum tangential displacement u_{xi} along the "stick" loading path at which a normal perturbation $^p u_{yj}$ produces "premature" sliding, when the driving velocity is $U_{xk}^{(0)}$.

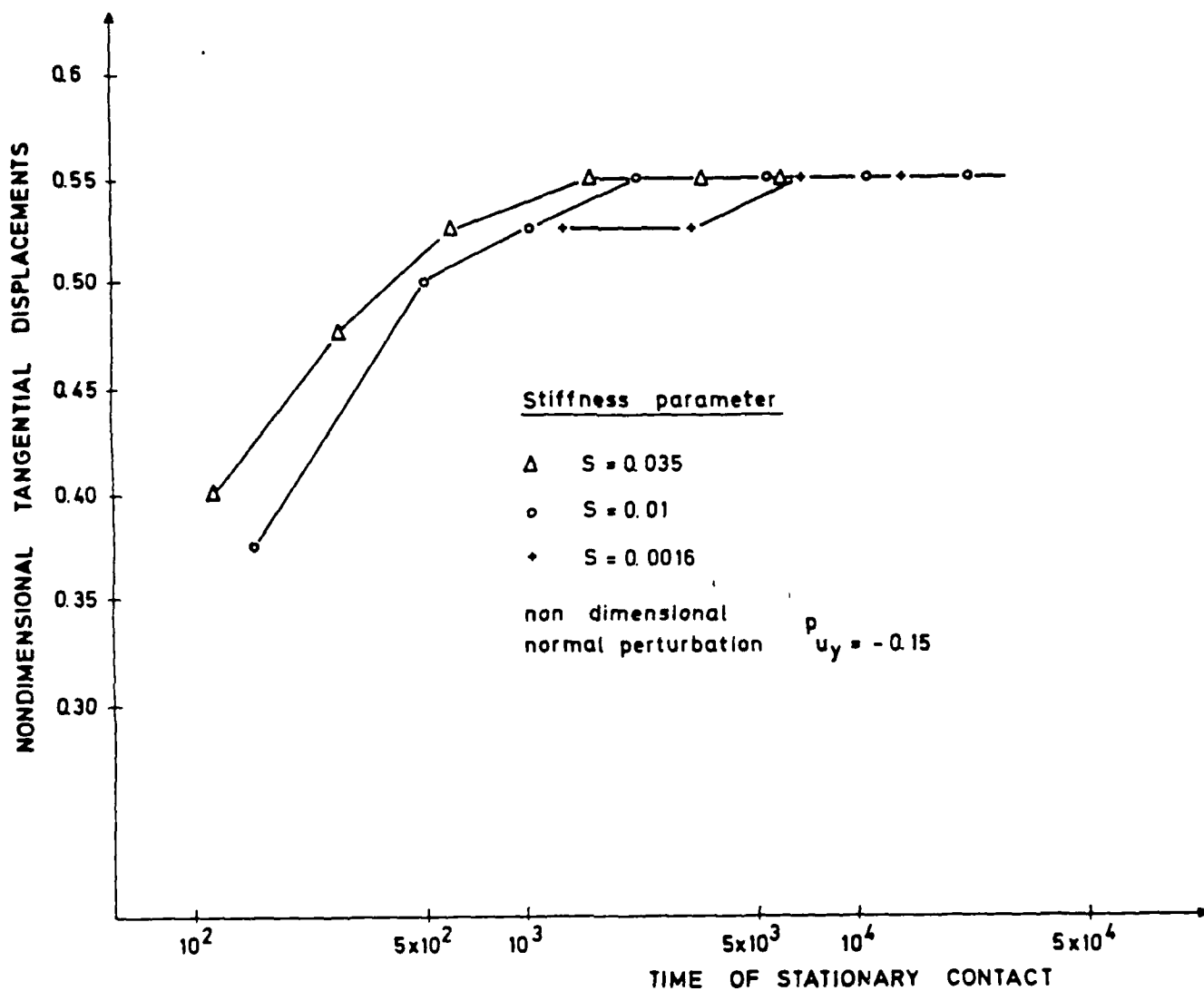


Figure 5.35: Minimum tangential displacement u_{xi} along the "stick" loading path at which a normal perturbation $p_{u_y} = 0.15$ produces "premature" sliding versus time of stationary contact (τ_{ki}/\sqrt{s}).

values of the time of stationary contact.

5.4.3 Discussion

The previous results, however suggestive they are, should be analyzed having in mind the admitted limitations of the model used: preliminary tangential displacements occurring before gross sliding and the details of the plastic deformation of the interface cannot be modelled with the constitutive laws (3.1,2) adopted in this work. Since these effects may have some importance along the quasistatic loading process studied here, the results above should be viewed only as indicators of what to expect when using more complex models of the interface behavior.

A question that is important to discuss is the size of the normal perturbations considered. As we decrease the size of the perturbation, its destabilizing effect of course is reduced (recall Fig. 5.34). It is easy to anticipate, and we confirmed it numerically, that sufficiently small normal perturbations have negligible effect on the initiation of sliding. However, we observe that the size of the perturbations considered here, although "mathematically not small", is indeed "physically quite small" and of the order of magnitude of the microseisms mentioned in Section 2.6. For static penetrations $Y = (W/c_n BL)^{1/m_n}$ in the range 0.3 to $10\mu\text{m}$, the perturbations considered here would be in the range 0.015 to $2\mu\text{m}$.

Apparent decreases of the coefficient of static friction with the increase of driving velocity, as suggested by the numerical results above, would affect the stick-slip results of Section 5.3.2 (recall that in the computations leading to Figs. 5.22 to 5.24 no reductions of static friction were taken into account). Smaller stick-slip amplitudes and more pronounced slopes of the amplitude-velocity curves in Figs. 5.22 and 5.23 should be expected as a result of the reductions of static friction. On the other hand, these reductions would lead to smaller periods of stick and consequently to larger frequencies in Fig. 5.24, closer to the natural frequency of the tangential motion of the system. These effects would certainly improve on what appears to be an "excessive" sharpness of the transition stick-slip to apparently smooth sliding in Figs. 5.22 and 5.23 and also on the smallness of the maximum stick-slip frequencies in Fig. 5.24.

It is thus clear that the "rate dependence" of the coefficient of static friction plays an important role on the stick-slip oscillations, precisely as it has been assumed in previous analyses. The question that we raise here and that needs further study is solely related with its origin — is it an intrinsic property of the contacting surfaces or is it the result of a (still not well defined) "instability" along the stick portion of the stick-slip cycles? Much

theoretical, numerical and experimental work is still needed before a definitive answer to this question can be given.

6 Critical Sliding of a Rigid Block On a Fixed Horizontal Plane

In this section, we study the sliding of a rigid rectangular-prismatic block on a fixed horizontal plane, when acted upon by a tangential force equal to the critical force of static friction. These conditions are essentially the same as those suggested by Euler more than two hundred years ago for an experiment on the difference between static and kinetic friction: a body sliding down an inclined plane at the critical slope at which sliding is initiated (recall Section 2.1).

6.1 Governing Equations

The equations of motion have the form (5.4) with M, P, Q , and J as in Section 5, but now we have $\dot{U}_x^C = 0$ (hence, $\dot{\Phi} = 0$), $K_x = C_x = 0$ (hence $K = C = 0$), and

$$F = \begin{Bmatrix} -F \\ W \\ 0 \end{Bmatrix}$$

where F is the absolute value of the applied tangential force (see Fig.6.1).

Since $K_x = 0$, the tangential displacement u_x does not appear in the governing system (5.4) and can be computed *a posteriori* from the tangential velocity and the initial condition \bar{u}_{x0} :

$$u_x(t) = \bar{u}_{x0} + \int_0^t \dot{u}_x(\sigma) d\sigma.$$

We can thus study the problem in the five-dimensional phase space of the displacements u_y and u_θ and the velocities v_x, v_y and v_θ . The governing system of five first-order differential equations or inclusions can be written in the form

$$\begin{aligned} \dot{u}^* &= v^* \\ M\dot{v}_x + J_x(u^*, v) &\ni -F \\ M^*\dot{v}^* + P^*(u^*) + Q^*(u^*, v^*) + J^*(u^*, v) &\ni F^* \end{aligned}$$

where we again use the superscript $*$ to denote subvectors or submatrices associated with normal and rotational degrees-of-freedom.

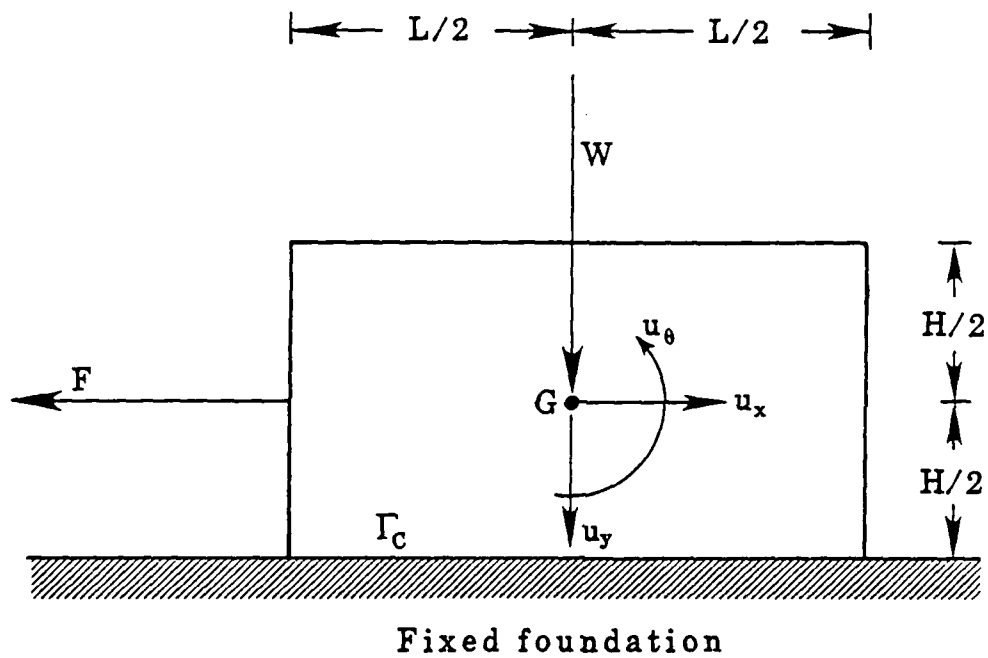


Figure 6.1: Geometry, degrees-of-freedom and forces applied to a prismatic block sliding with friction on a fixed foundation.

The singular points $[u_{y0}, u_{\Theta 0}, v_{x0}, v_{y0}, v_{\Theta 0}]$ for this system, if they exist, satisfy the conditions

$$\mathbf{v}_0^* = 0 \quad (6.1)$$

$$J_x(\mathbf{u}_0^*, \mathbf{v}_0) \ni -F \quad (6.2)$$

$$\mathbf{P}^*(\mathbf{u}_0^*) + \mathbf{J}^*(\mathbf{u}_0^*, \mathbf{v}_0) \ni \mathbf{F}^* \quad (6.3)$$

where, due to (6.1), $\mathbf{v}_0 = [v_{x0}, 0, 0]^t$. These singular points correspond to steady sliding motions of the block towards the left when $\text{sgn}(v_{x0}) = -1$. For that situation (6.2) and (6.3) reduce to

$$J_{\eta x}(\mathbf{u}_0^*) = -F \quad (6.4)$$

$$\mathbf{P}^*(\mathbf{u}_0^*) + \mathbf{J}_{\eta}^*(\mathbf{u}_0^*) = \mathbf{F}^* \quad (6.5)$$

where \mathbf{J}_{η} was defined in (5.7), and now

$$\eta \stackrel{\text{def}}{=} -\text{sgn}(v_{x0}) = +1.$$

Equations (6.5) determine the normal and rotational displacements for the non-isolated (because $v_{x0} \leq 0$ is arbitrary) steady-sliding singular points of the system. Equation (6.4) determines the force F required for such a steady-sliding.

We emphasize that, since no distinctions between static and kinetic friction were considered, when F has the critical value (6.4), steady-sliding at an arbitrarily small speed (including the limit case $v_{x0} = 0$) is possible, in apparent contradiction with Euler's remarks.

However, we can observe that the two equations of the system (6.5) are precisely the same as the second and third equations of the system (5.6). Furthermore, for any *strictly negative* v_{x0} and denoting $N_{\delta}(\mathbf{v}_0) = \{\mathbf{v} \in \mathbf{R}^3 : |v_x - v_{x0}| + |v_y| + |v_{\Theta}| < \delta = |v_{x0}| \max(1, H/2)\}$ the governing system reduces to

$$M\dot{v}_x + J_{\eta x}(\mathbf{u}^*) = -F \quad (6.6)$$

$$\left. \begin{aligned} \dot{\mathbf{u}}^* &= \mathbf{v}^* \\ \mathbf{M}^* \dot{\mathbf{v}}^* + \mathbf{P}^*(\mathbf{u}^*) + \mathbf{Q}^*(\mathbf{u}^*, \mathbf{v}^*) + \mathbf{J}_\eta^*(\mathbf{u}^*) &= \mathbf{F}^* \end{aligned} \right\} \quad (6.7)$$

for every $(\mathbf{u}^*(t), \mathbf{v}(t)) \in \mathbf{R}^2 \times N_\delta(\mathbf{v}_0)$. It is clear that the system of equations (6.7) is equivalent to the second and third equations in (5.8). Linear stability analysis of the normal and rotational motions leads here again to the eigenvalue problem (5.25). This means that *all the results of Section 5.2 on the steady-sliding normal and rotational displacements (u_{y0} and $u_{\theta 0}$) as well as on the linear stability of normal and rotational motions in the neighborhood of steady-sliding apply thus to the present problem.*

Before presenting numerical results for this problem, we still need to indicate the nondimensional variables used. The nondimensional displacements and time are defined again by (5.16) and (5.29), respectively. Y in (5.16) is again given by (5.18) but now the choices (5.17) and (5.30) for X and ω in (5.16) and (5.29) are not possible since $K_x = 0$. We adopt here $\omega = \omega_{y0}$ with ω_{y0} defined by (5.27) and

$$X = Y/m_n \quad (6.8)$$

With these variables, the nondimensional governing system is again of the form (5.31) with $\mathbf{K} = \mathbf{C} = \mathbf{0}$, $\Phi' = \mathbf{0}$ and \mathbf{M} , \mathbf{P} , \mathbf{Q} , \mathbf{J} as in Section 5.3.1 but with m_n replacing the parameter s in all its occurrences. The nondimensional vector of applied forces is now

$$\mathbf{F} = \begin{Bmatrix} -F/W \\ m_n \\ 0 \end{Bmatrix}$$

6.2 Numerical Results

Similarly to Section 5.3, interesting situations arise when the governing parameters (h , f , $\hat{\varepsilon}$, m_n , m_T and l_n) are chosen so that the steady-sliding of the block is unstable.

For each set of parameter values we perform a transient analysis, starting with a small sliding velocity towards the left ($\bar{u}_{x1} = v_{x0} < 0$) and with normal and rotational displacements and velocities (\bar{u}_{y0} , $\bar{u}_{\theta 0}$, \bar{u}_{y1} , $\bar{u}_{\theta 1}$) assuming values that result from the superposition of a small perturbation (${}^p u_y$, ${}^p u_\theta$, ${}^p v_y$, ${}^p v_\theta$) on the steady-sliding state (u_{y0} , $u_{\theta 0}$, $v_{y0} = 0$, $v_{\theta 0} = 0$). Arbitrarily, the initial tangential displacement (\bar{u}_{x0}) is assumed to be zero. The initial conditions are thus

$$\bar{u}_{x0} = 0$$

$$\bar{u}_{y0} = u_{y0} +^p u_y$$

$$\bar{u}_{\Theta 0} = u_{\Theta 0} +^p u_{\Theta}$$

$$\bar{u}_{x1} = v_{x0} < 0$$

$$\bar{u}_{y1} =^p v_y$$

$$\bar{u}_{\Theta 1} =^p v_{\Theta}$$

In a manner similar to the study in Section 5.3, the instability of the steady-sliding originates a growing normal and rotational oscillation (Figures 6.2 to 6.5) which produces a variation of the maximum instantaneously available friction force. As a result, the body advances intermittently with successive periods of stick and slip. The average contact friction force for that intermittent motion is smaller than the constant applied force F so that, in average, the body accelerates as shown in Figures 6.6 to 6.8. As the average tangential speed of the body increases two apparently distinct types of behavior were observed:

1. for values of the friction parameter f smaller than some critical value, dependent on the other parameters involved, and after the body attains a sufficiently large average speed, no more stick states are possible (see Figs. 6.9 and 6.10) and the normal and rotational oscillations cease to increase and remain bounded (Figures 6.2 to 6.4) while the body advances with an average velocity that is much larger than its initial value (Figures 6.6 to 6.8)
2. for values of f greater than such a critical value, the normal and rotational oscillations keep growing (Fig. 6.5) and the body advances in an irregular manner, jumping and tumbling; since periods of stick continue to occur whenever contact is established (Fig. 6.11), the body keeps accelerating for all the time spanned by our analysis (Fig. 6.6 for $f=0.9$ or 1.2).

A precise mathematical characterization of the behaviors described above is not yet available. We remark here only that, in case (1) above, the complexity of the observed trajectories, the "size" and "thickness" of the regions (in the four-dimensional space of the variables $u_y, u_{\Theta}, v_y, v_{\Theta}$) to which the trajectories are attracted depend on the values assumed for the various parameters. Consequences of this on the tangential motion are illustrated in Fig. 6.6 (the larger the friction parameter f , the larger the sliding velocity attained) and in Fig. 6.7 (the larger the normal damping parameter \hat{z} , the smaller the sliding velocity attained). In Fig. 6.8 we show that the final average sliding velocity attained in case (1) is not affected by changing the small initial sliding velocity or the small perturbations to the

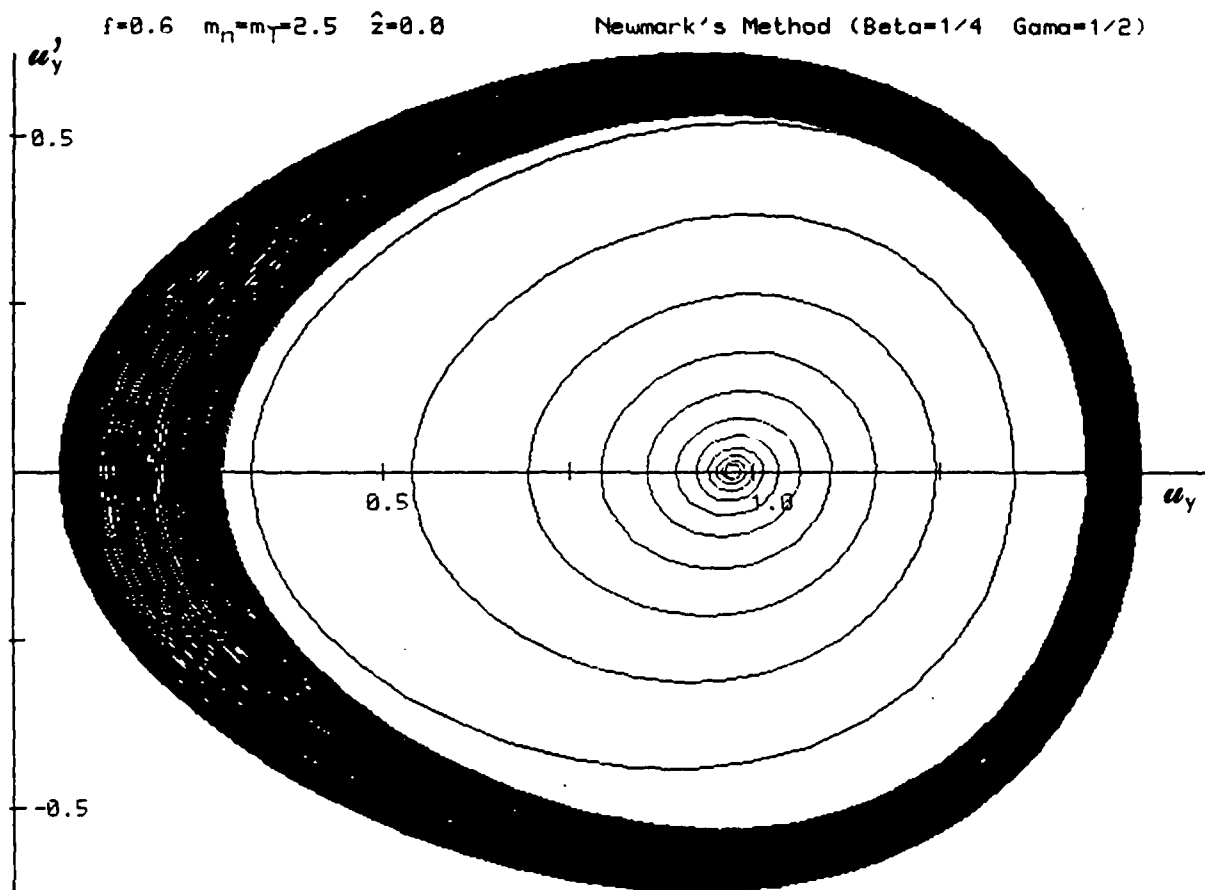


Figure 6.2: Phase plane plots of the normal oscillation of a rigid block acted by a tangential force equal to the critical force of static friction obtained with Newmark's method. $f = 0.6$; $h = 0.45$; $m_n = m_T = 2.5$; $\hat{z} = 0.0$. Motion of the type (1) in the text.

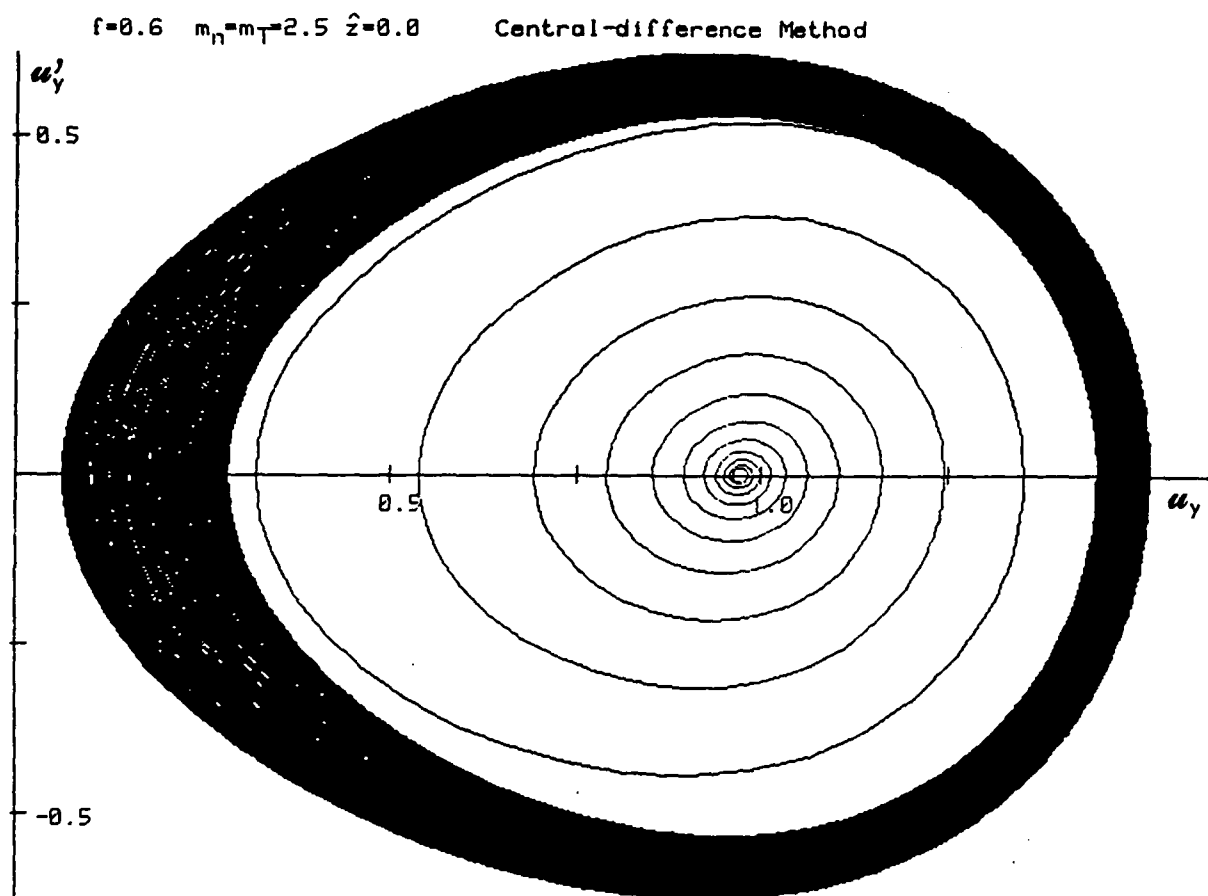


Figure 6.3: Phase plane plots of the normal oscillation of a rigid block acted by a tangential force equal to the critical force of static friction obtained with the central difference method. $f = 0.6$; $h = 0.45$; $m_n = m_T = 2.5$; $\hat{z} = 0.0$. Motion of the type (1) in the text.

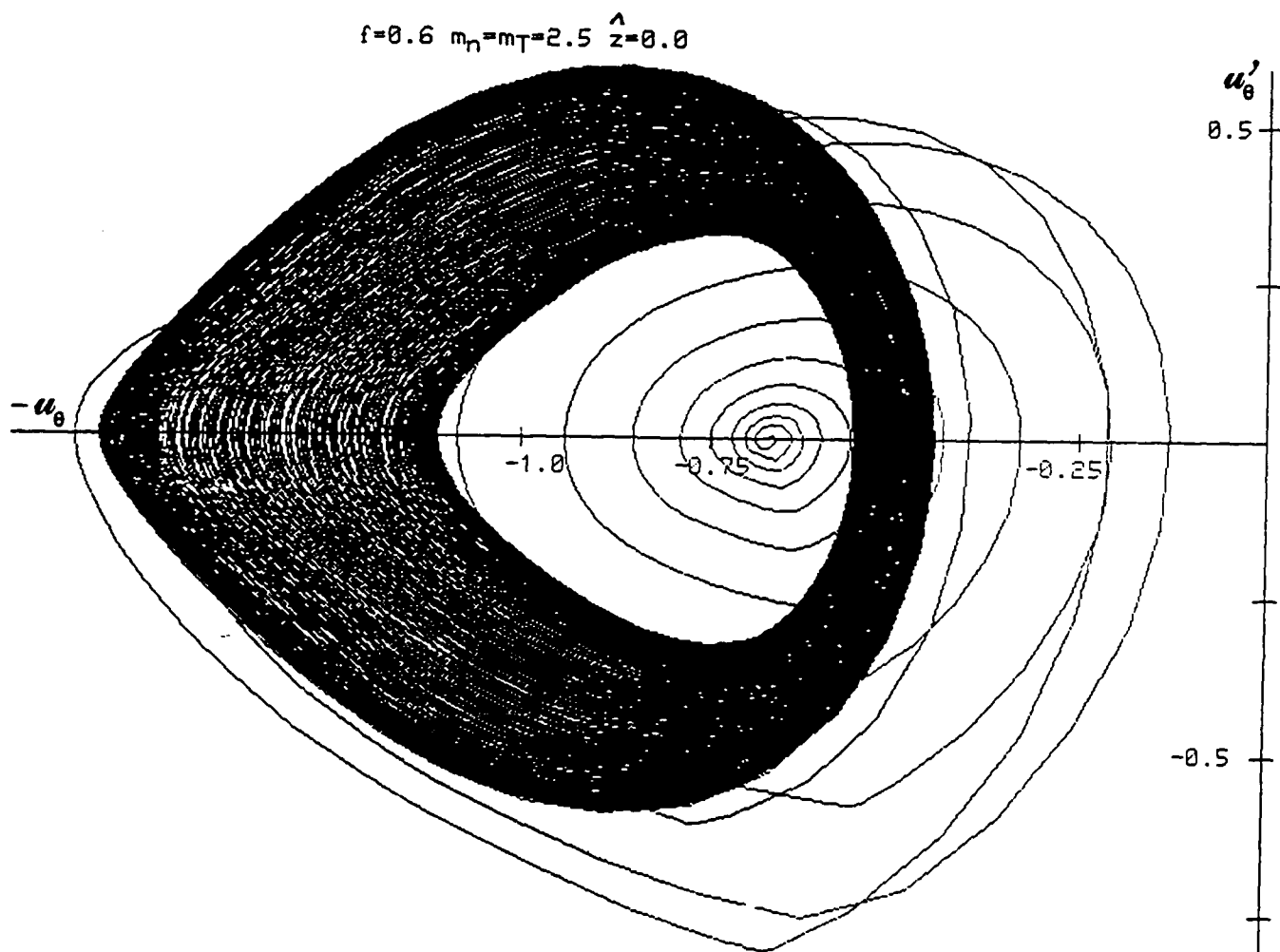


Figure 6.4: Phase plane plot of the rotational oscillation of a rigid block acted by a tangential force equal to the critical force of static friction. $f = 0.6$; $h = 0.45$; $m_n = m_T = 2.5$; $\hat{z} = 0.0$. Motion of the type (1) in the text.

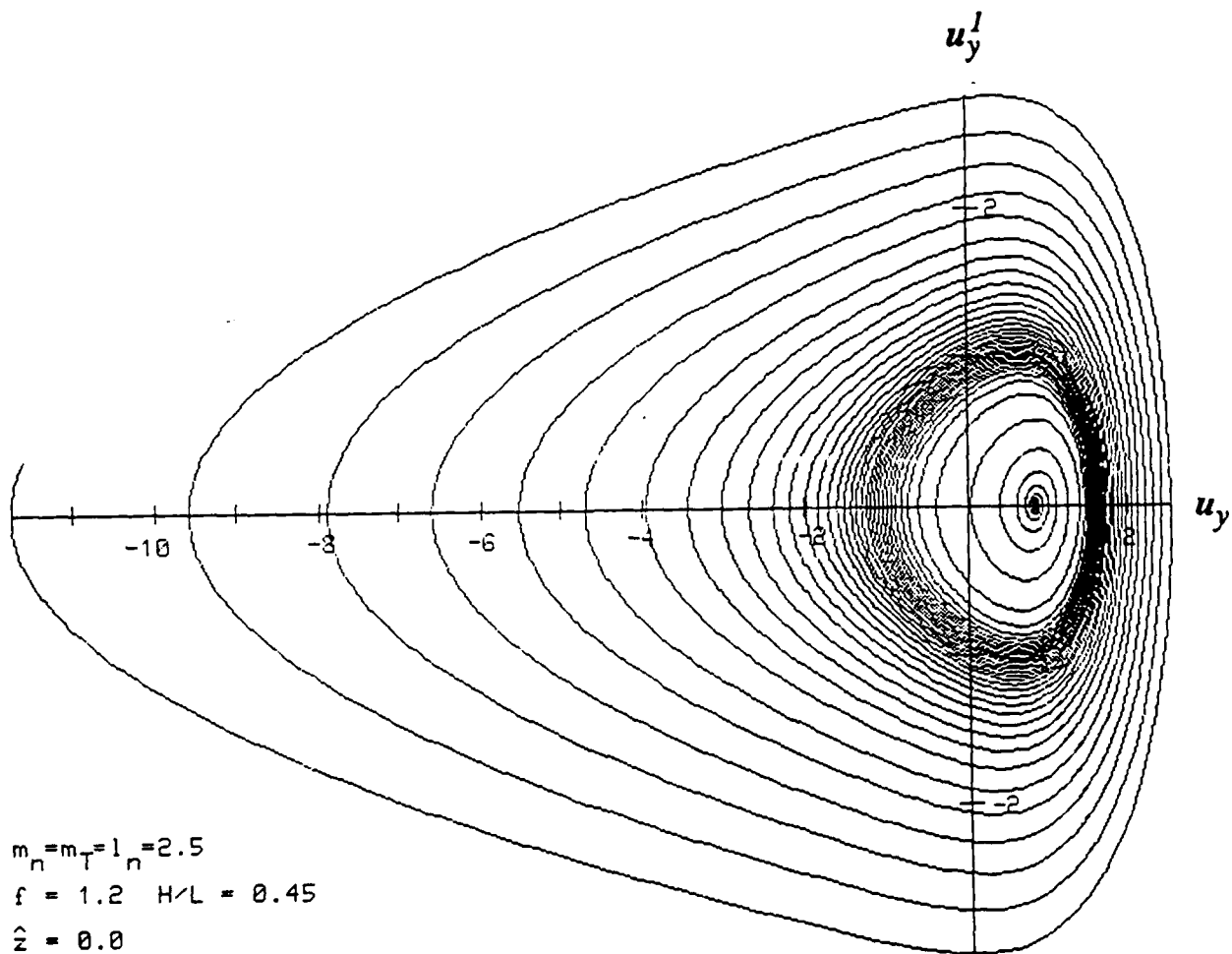


Figure 6.5: Phase plane plot of the initial part of the normal oscillations of a rigid block acted by a tangential force equal to the critical force of static friction, for a large value of the friction parameter $f = 1.2$; $h = 0.45$; $m_n = m_T = 2.5$; $\hat{z} = 0.0$. Motion of type (2) in the text.

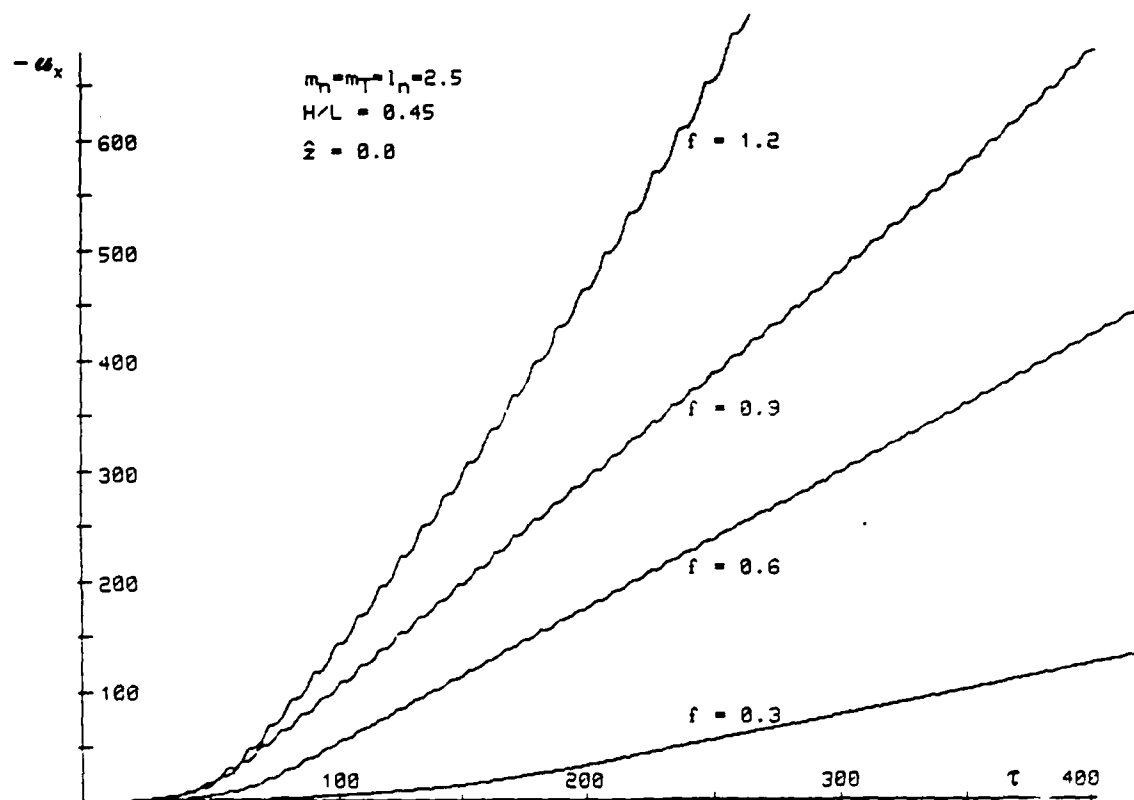


Figure 6.6: Effect of the value of the friction parameter on the tangential motion of a rigid block acted by a tangential force equal to the critical force of static friction. $h = 0.45$; $m_n = m_T = 2.5$; $\hat{z} = 0.0$. For $f = 0.3, 0.6$ the motion is of type (1) and for $f = 0.9, 1.2$ the motion is of type (2).

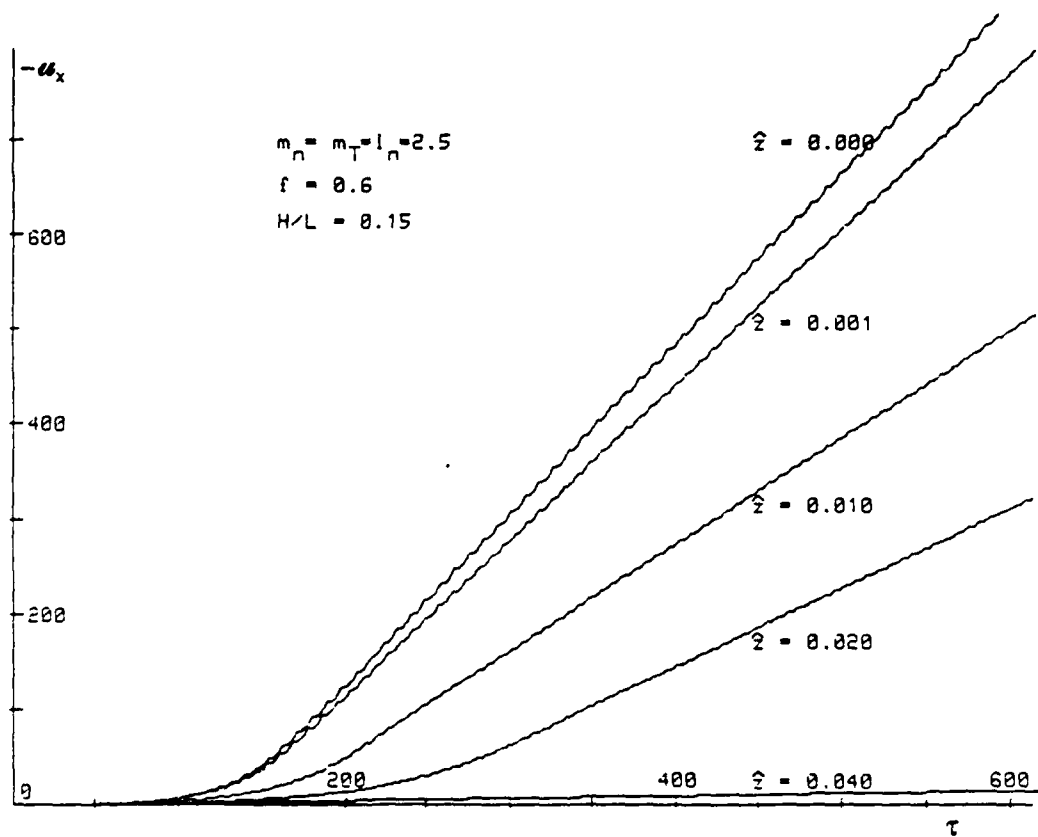


Figure 6.7: Effect of the value of the normal interface damping parameter (\hat{z}) on the tangential motion of a rigid block acted by a tangential force equal to the critical force of static friction. $h = 0.15$; $m_n = m_T = l_n = 2.5$; $f = 0.6$. All the motions are of type (1).

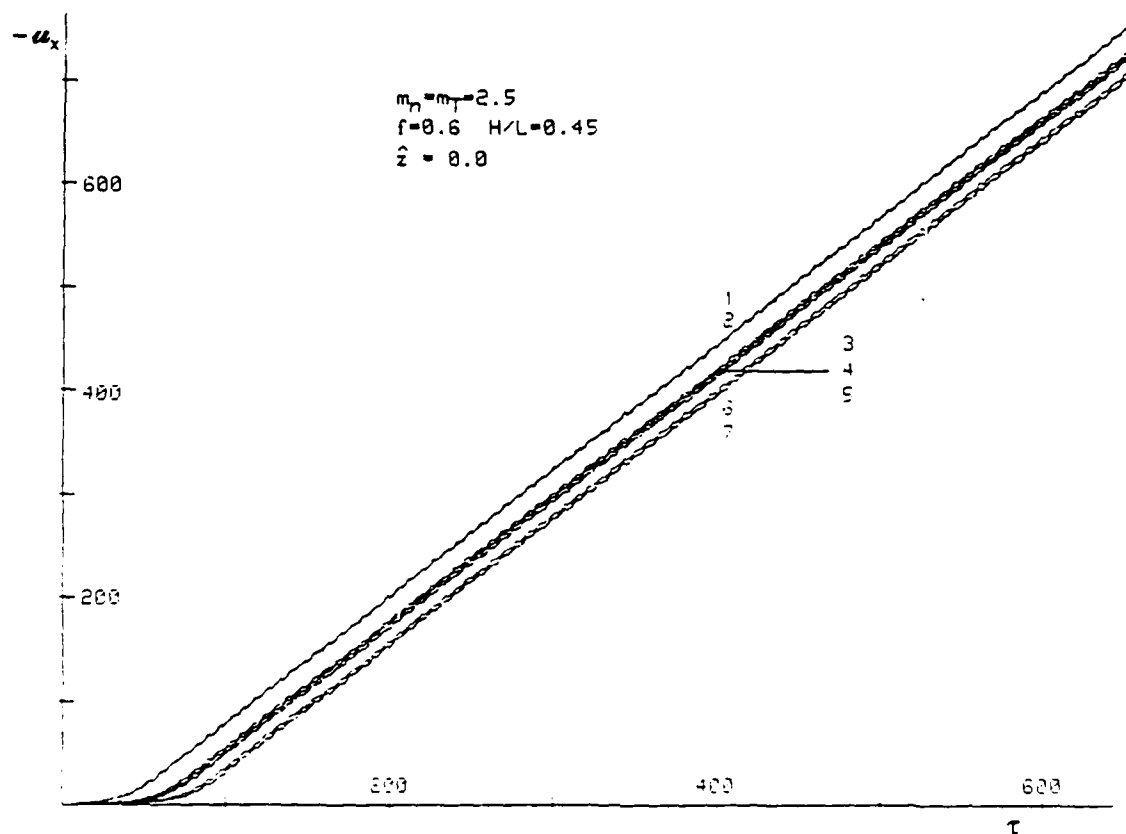


Figure 6.8: Effect of various initial conditions on the final average velocity for a motion of type (1).

	$h = 0.45,$	$m_n = m_T = 2.5,$	$f = 0.6; \hat{z} = 0.0$
1	$v_{x0} = -0.01,$	$v_{y0} = v_{\theta 0} = -0.01,$	$v_{u_y} = v_{u_\theta} = 0$
2	$v_{x0} = -0.01,$	$v_{y0} = -0.01,$	$v_{u_y} = v_{u_\theta} = v_{v_\theta} = 0$
3	$v_{x0} = -0.01,$	$v_{y0} = -0.01,$	$v_{u_y} = v_{u_\theta} = v_{v_\theta} = 0$
4	$v_{x0} = -0.01,$	$v_{y0} = -0.01,$	$v_{u_y} = v_{u_\theta} = v_{v_\theta} = 0$
5	$v_{x0} = -0.01,$	$v_{y0} = v_{\theta 0} = -0.01,$	$v_{u_y} = v_{u_\theta} = 0$
6	$v_{x0} = -0.01,$	$v_{y0} = 0.01,$	$v_{u_y} = v_{u_\theta} = v_{v_\theta} = 0$
7	$v_{x0} = -0.01,$	$v_{y0} = -0.01,$	$v_{u_y} = v_{u_\theta} = v_{v_\theta} = 0$

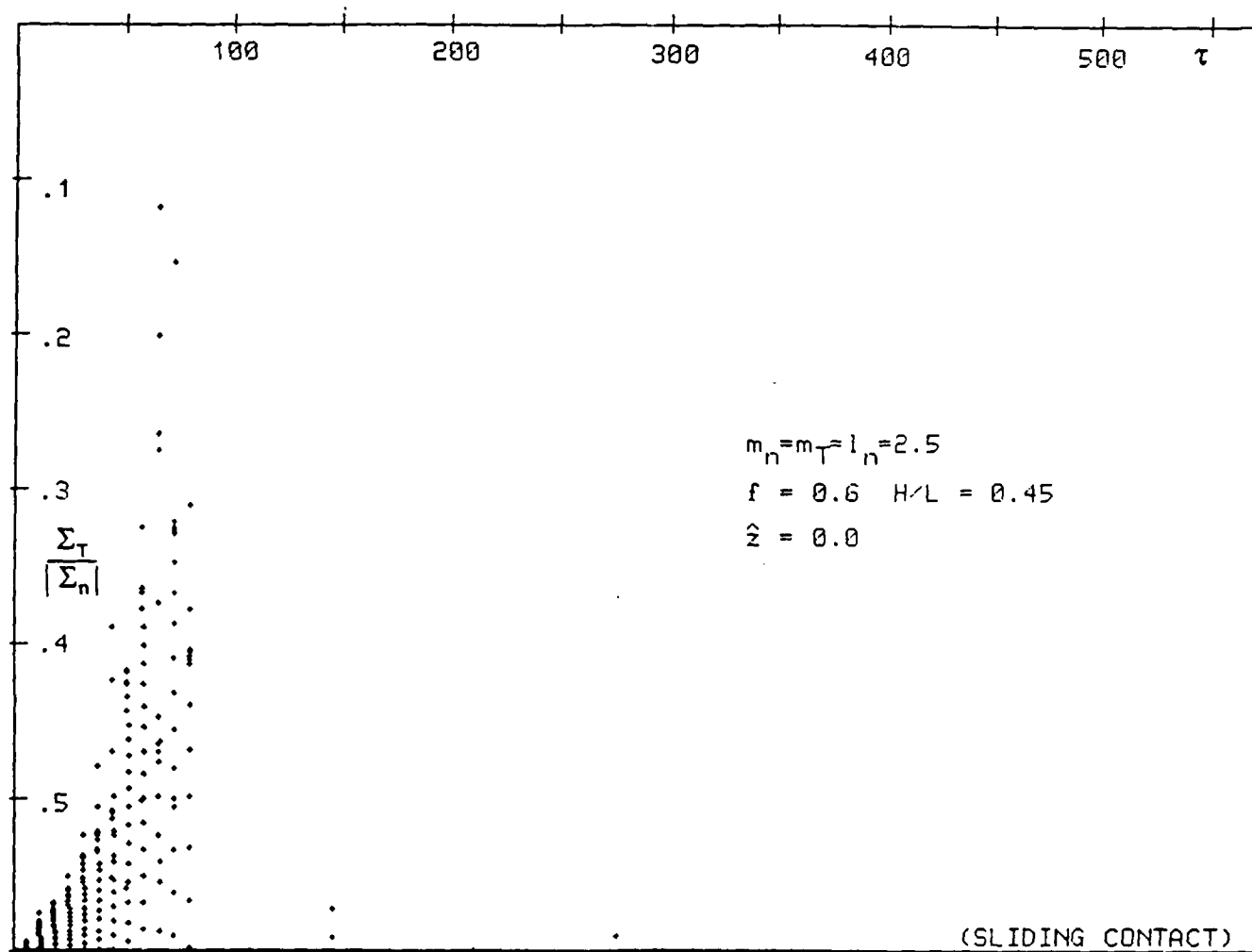


Figure 6.9: Computed values of the ratio: instantaneous friction force/instantaneous normal force ($\Sigma_{Tx}(t)/|\Sigma_{ny}(t)|$). $m_n = m_T = 2.5$; $h = 0.45$; $f = 0.6$; $\hat{z} = 0$. Motion of the type (1) in the text. After the initial accelerating phase the body acquires a sliding motion with continuous contact. Not all the computed values are plotted and, for clarity, no line connects the successive plotted values.

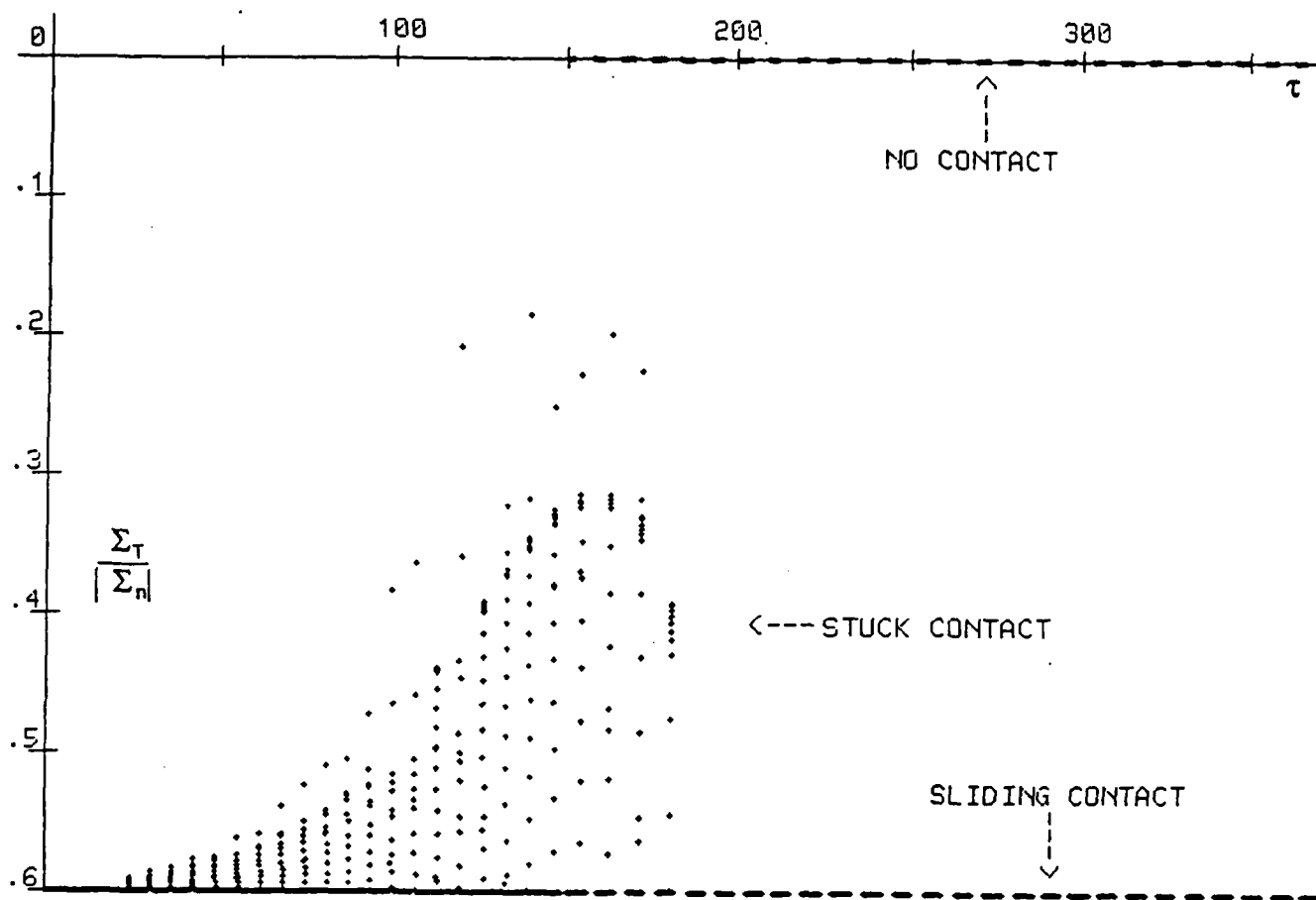


Figure 6.10: Computed values of the ratio: instantaneous friction force/instantaneous normal force ($\Sigma_{Tx}(t) / |\Sigma_{ny}(t)|$). $m_n = m_T = 2.5$; $h = 0.15$; $f = 0.6$; $\hat{z} = 0$. Motion of the type (1) in the text. After the initial accelerating phase the body advances with successive periods of contact and no contact, but sliding occurs whenever contact is established. Not all the computed values are plotted; for clarity, no line connects the successive plotted values; for no contact situations, the value zero is arbitrarily assigned to the ratio $\Sigma_{Tx}/|\Sigma_{ny}|$.

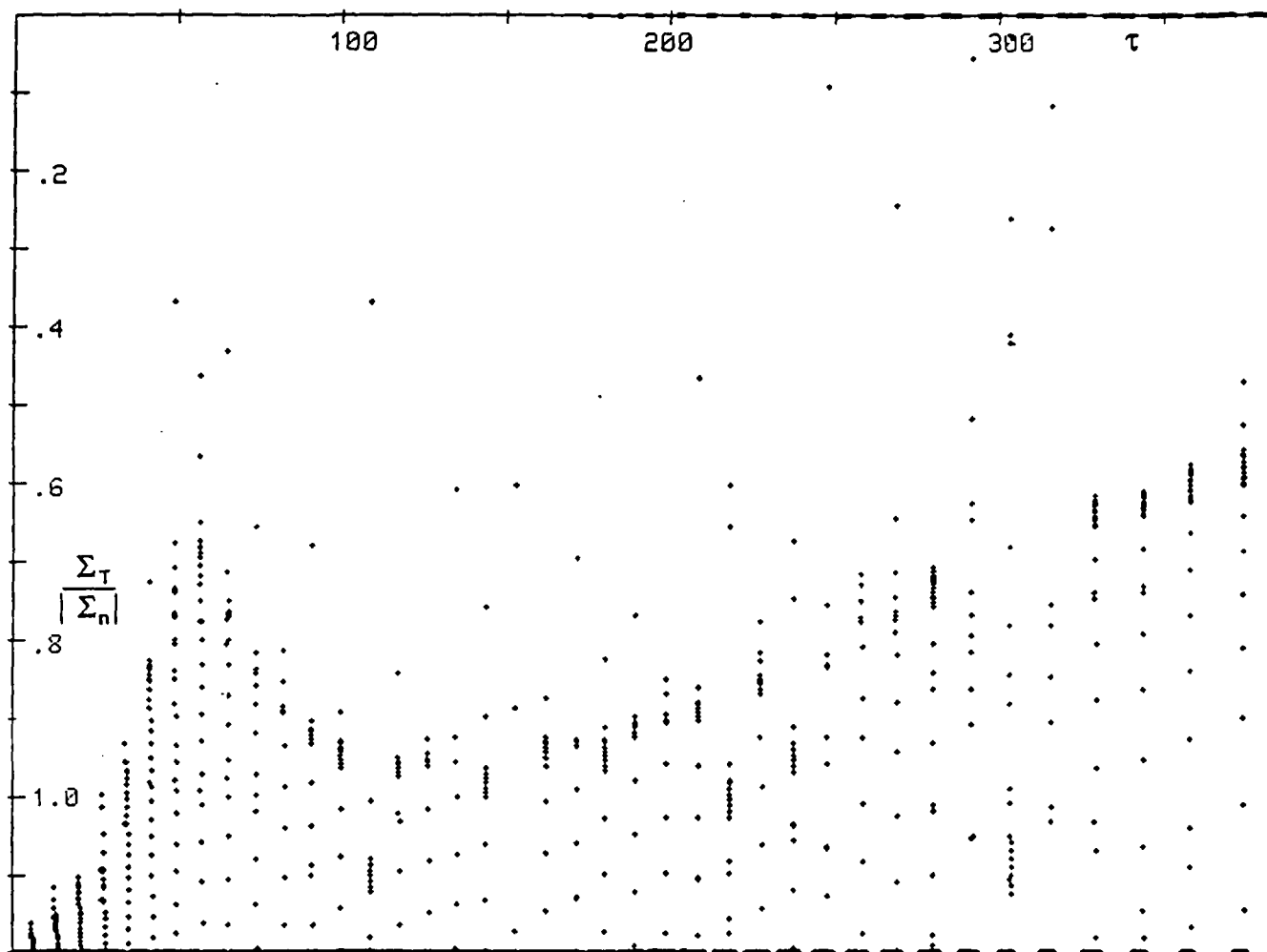


Figure 6.11: Computed values of the ratio: instantaneous friction force/instantaneous normal force ($\Sigma_{Tx}(t)/|\Sigma_{ny}(t)|$). $m_n = m_T = 2.5$; $h = 0.45$; $f = 1.2$; $\hat{z} = 0$. Motion of the type (2) in the text. For all the time spanned by the analysis the body keeps accelerating; periods of stick exist whenever contact is established. Not all the computed values are plotted; for clarity no line connects the successive plotted values; for no contact situations the value zero is arbitrarily assigned to the ratio $\Sigma_{Tx}/|\Sigma_{ny}|$.

steady sliding state. Finally, we observe that during the final part of the motions of the type (1) the observable average coefficient of kinetic friction is equal to the coefficient of static friction (no thermal softening effects are taken into account in the present analysis).

With respect to case (2) we remark only that, as the normal and rotational oscillations increase, the problem becomes less and less a problem of sliding friction and becomes more and more a problem of impact with friction, for which neither the physical assumptions (e.g. the normal interface law) nor the geometric assumptions (small rotations) used here are appropriate. Of course, sufficient normal interface damping (\hat{z}) may transform, with all the other parameters remaining the same, a behavior of the type (2) into a behavior of the type (1): see Figs. 6.12 and 6.13.

6.3 Discussion

Although much work has to be done before a full picture of the qualitative behavior summarized above is achieved, the essential point to be retained from the previous examples is that *a rigid block, when acted upon by a tangential force equal to the critical force of static friction, may have an accelerated sliding even when the coefficients of static and kinetic friction are equal.* The same can be said also about the experiment proposed by Euler [33] with a block on an inclined plane at the critical slope.

We need to have an idea of the order of magnitude of the sliding velocities acquired exclusively as a consequence of the effects studied above. Of course, that magnitude depends much on the geometries of the body and the contact. For the geometry and data in Figures 6.6 to 6.8 the nondimensional velocities u'_x are of the order 1 for motions of the type (1). Since

$$\dot{u}_x = \sqrt{gY/m_n} u'_x$$

we have, for $g \approx 10\text{ms}^{-2}$, $m_n \in [2, 3]$ and $Y \in [0.3, 10]\mu\text{m}$, sliding velocities in the range 1 to 7 mm/s. This means that motions of the type (1) can explain initial “jerks” experimentally observed on the transition from repose to sliding. We observe in this respect that, in their study on friction at low speeds, Jenkin and Ewing [44] considered a requirement for the selection of their experimental apparatus that *“the change should be from motion to rest rather than from rest to motion, so as to avoid any jerk at the instant of passage from one to the other state.”* The importance of this consideration on the success of their experimental work becomes clear if we notice the small sliding velocities they were able to use (0.06 to 3 mm/s) without the misleading effect of initial “jerks”. On the other hand, motions of the

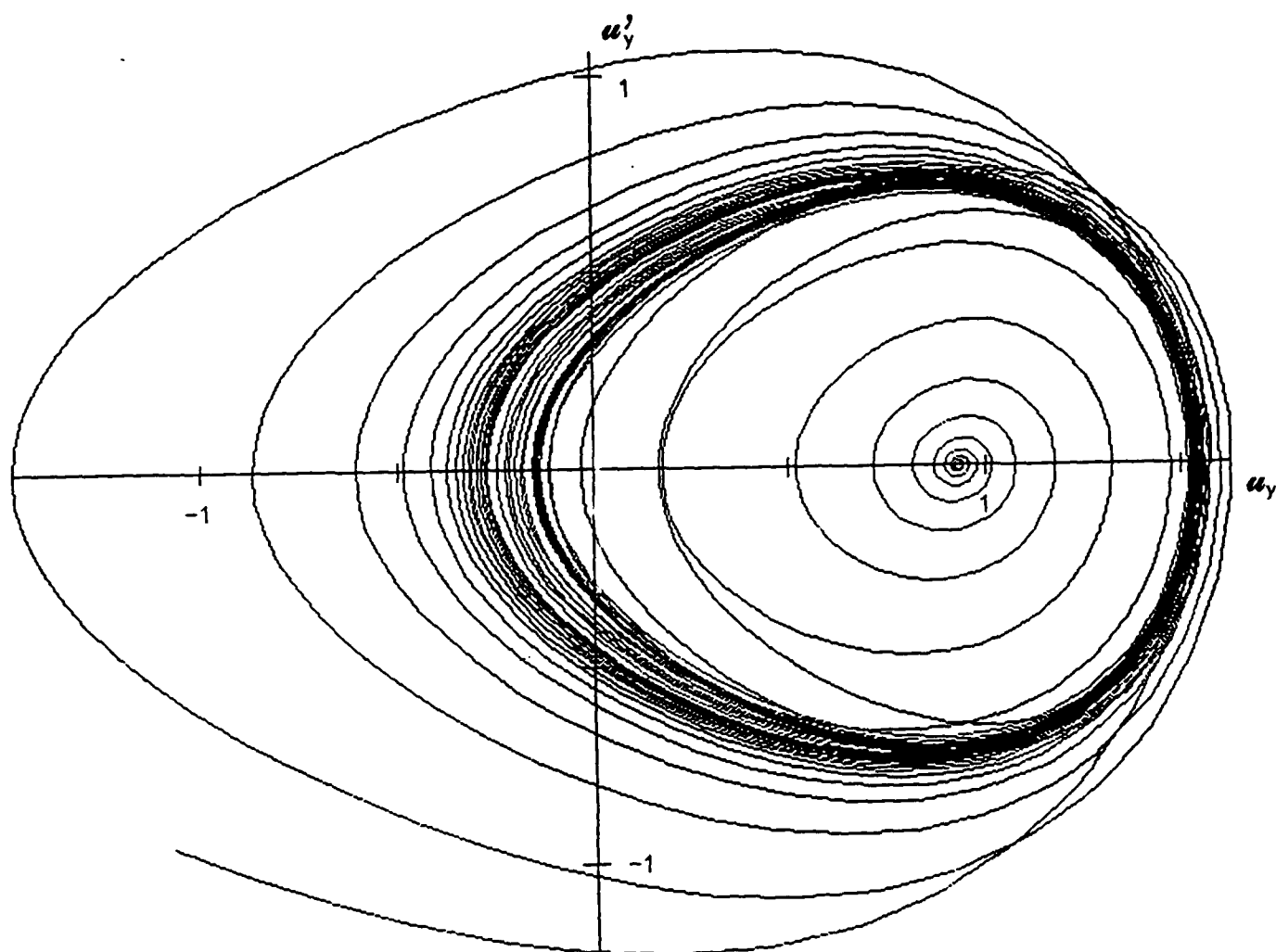


Figure 6.12: Effect of the normal interface damping on the resulting type of motion. Phase plane plots of the normal oscillation. $m_n = m_T = l_n = 2.5$; $h = 0.45$; $f = 0.9$. $\dot{z} = 0.0$: motion of type (2)

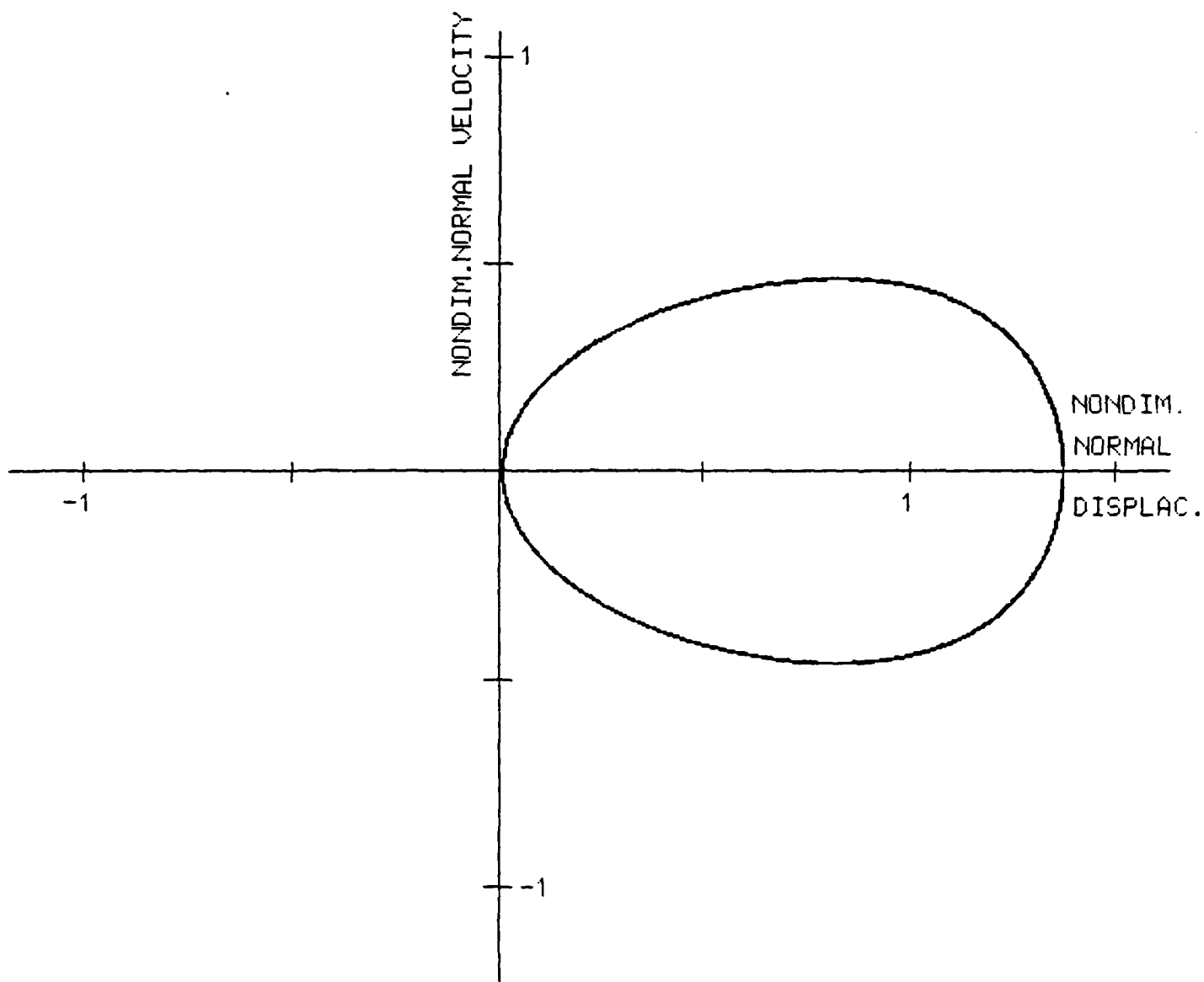


Figure 6.13: Effect of the normal interface damping on the resulting type of motion. Phase plane plots of the normal oscillation. $m_n = m_T = l_n = 2.5$; $h = 0.45$; $f = 0.9$. $\dot{z} = 0.005$; final portion of a motion of type (1)

type (2), together with thermal softening effects after sufficiently large sliding speeds are attained, can probably explain more prolonged accelerations of the sliding body.

7 Conclusions

1. Experimentally observed dependences of the coefficient of static friction on the time of stationary contact, distinctions between coefficient of static and kinetic friction and dependences of the latter on the sliding speed at small sliding speeds are not necessarily intrinsic properties of dry metallic interfaces. Physically small external perturbations and the geometric and dynamic properties of an experimental apparatus may lead to the experimental observation of these phenomena even when the coefficient of friction is a constant independent of sliding speed or time of contact.
2. For sufficiently large coefficients of friction and appropriate values of the other governing parameters, steady-sliding of a block-on-slideway system is dynamically unstable. That instability results from the inherent nonsymmetry of the frictional contributions to the governing equations and it may occur even when the coefficient of friction does not depend on the sliding speed.
3. Instability of steady-sliding leads to normal oscillations, the frequency of which is governed by the interface normal stiffness properties and the dynamic properties of the experimental apparatus. The corresponding oscillations of the maximum instantaneously available friction force originate a high-frequency stick-slip motion which in turn leads to apparent coefficients of kinetic friction lower than the coefficient of static friction, precisely as proposed originally by Budanov, Kudinov and Tolstoi.
4. Also, as proposed by the same authors, increases of apparent coefficient of kinetic friction with increasing sliding speeds may be the result of a decrease of the periods of stick of those high-frequency motions when the average sliding speed increases.
5. The same high-frequency stick-slip motions may also be associated with decreases of apparent kinetic friction with the increase of sliding speed. This is essentially the result of an increase of the amplitudes of those high-frequency self-excited oscillations when the driving speed is increased. For some ranges of the governing parameters the high-frequency mechanism (mechanism II in Section 2.6) may thus lead to effects of the

type attributed by Tolstoi and coworkers to the assymetry of the normal oscillations mechanism (mechanism I in Section 2.6).

6. The acceleration of a body down an inclined plane at the critical slope at which sliding initiates is not necessarily the result of a drop of the coefficient of friction upon the initiation of sliding. It may be solely the result of the dynamic instability of the steady-sliding descent of the body and the consequent normal (and rotational) oscillations.
7. Experimentally observed loops described by the coefficient of friction during the slip phase of stick-slip motions are essentially meaningless if corresponding changes on normal contact force are not appropriately taken into account. Different shapes of those friction-velocity loops may simply result from different forms of the normal jumps that accompany the sliding portion of those stick-slip motions.

CHAPTER 3

NUMERICAL STUDY OF FRICTIONAL BEHAVIOR OF SLIDING SYSTEMS

1 Introduction

Contact and friction phenomena have been of interest to researchers since the historical works of Amontons [1699] and Coulomb [1785]. Experimental results on friction and corresponding theories developed before the early 1960's are summarized in the books of Bowden and Tabor [1950,1964]. As a general rule, the explanation of all aspects of frictional behavior has been attributed to specific properties of the contacting surfaces. In particular, the difference between static and kinetic friction was assumed to be an intrinsic property of the frictional interface.

It was not until recently that the importance of the dynamic characteristics of the testing apparatus in the results of frictional experiments was explicitly noted. Although discrepancies between results of frictional experiments performed on different apparatus were observed by various researchers (see, e.g., Burwell and Rabinowicz [1953], Bodner and Tabor [1964]), the first clear statement of importance of dynamic characteristics and vibrations of the apparatus is due to Tolstoi [1967] and Tolstoi, Borisova and Grigorova [1971]. These researchers analyzed experimentally kinetic friction in the presence (and absence) of vibrations. They concluded that the *interface* coefficient of friction does not explicitly depend on sliding velocity, and the difference between the *apparent* static and kinetic coefficients of friction is the consequence of microscopic vibrations, inevitably accompanying frictional sliding.

This observation was confirmed by experiments of other researchers, investigating the influence of vibrations on the static and kinetic coefficients of friction—Weic [1962], Pohlman and Lehffeldt [1965], Godfrey [1967] or more recently Broniec and Lenkiewicz [1980], Aronov, D'Souza, Kalpakjian and Shareef [1984] and Chiou, Kato and Abe' [1987].

Observations and ideas presented by Tolstoi led Oden and Martins [1985] to a new approach to analysis of dynamic friction. They considered a relatively simple constitutive model of the interface, with a power law normal response and the coefficient of friction independent of velocity. This model was combined with an analysis of motion, in particular of normal vibrations, of the slider, to give apparent kinetic coefficient of friction different—in general—from the interface coefficient of friction. In their numerical studies Oden and Martins [1985] obtained a good qualitative modeling of general experimental observations. Some of their recent results are presented elsewhere in this report (Chapter 2).

The above approach was later studied by Becker and Tworzydło [1988] who applied it to

numerical modeling of reduction of static friction by vibrations. The obtained results were in very good agreement with experimental observations of Tolstoy [1966].

In this chapter we investigate further possibilities of application of the Oden/Martins model to realistic modeling of phenomena of dynamic friction. Our analysis focuses on the typical pin-on-disk experimental apparatus, which is a good representative of a broad class of mechanical sliding systems. In particular, we study self-excited oscillations of the slider, identify basic mechanisms leading to these oscillations as well as parameters of the system affecting their occurrence.

The general objective is to obtain both qualitative and quantitative agreement between our numerical predictions and experimental observations. As a particular object of our computations we have selected the representative pin-on-disk apparatus, investigated experimentally by Aronov, D'Souza, Kalpakjian and Shareef [1984].

The chapter is divided into two general parts. In the first part (Sections 2 through 5) we present a brief discussion of the most important phenomena, affecting the nature of dynamic friction and the value of kinetic coefficient of friction. We also present formulation of a boundary-value problem for two elastic bodies in sliding contact. A specially defined reference frame is devised to represent the motion of the sliding deformable surface (e.g., disk). A variational formulation of the boundary-value problem and two versions of its discretization are presented.

In the second part of this chapter (Sections 6 through 10) we present a study of frictional behavior of a pin-on-disk apparatus. We identify the basic mechanism of occurrence of self-excited oscillations and analyze the most important parameters of the system which affect this occurrence (stiffness, damping, angle of attack). Then, by means of transient analysis, we study the motion of the system in the self-excited zone and the resulting reduction of the kinetic coefficient of friction.

The theory, numerical approach and observations presented in this report can be generalized to a variety of mechanical systems with friction. Still, further research is necessary in order to enable more general practical applications. The most important directions of this research are identified in the last section of the report.

Part I

Formulation

2 Remarks on the Nature of Dynamic Friction

In this section we will review the most important phenomena, which affect the nature of dynamic friction, various forms of frictional sliding and the value of the kinetic coefficient of friction.

In the spirit of our remarks on importance of the overall dynamic characteristics of the sliding systems, we will classify each analyzed phenomena into one of the two groups:

- A. phenomena resulting primarily from properties of the interface, and
- B. phenomena associated with the dynamic characteristic of the system as a whole.

Our review will be based primarily on representative experimental works and numerical experiments of Oden and Martins [1985], continued and extended by Becker and Tworzydło [1988]. Some remarks will inevitably anticipate results presented in Part II of this paper. All of the phenomena considered in this section are listed in Table 1, which also contains a brief description and evaluation of each.

2.1 Dynamic Instability of the System

Dynamic instability of the system is a typical phenomenon of group B. The general idea is that if the system under consideration is dynamically unstable, then any perturbation of the so-called steady-sliding equilibrium position causes propagation of oscillations and occurrence of high amplitude vibrations of the systems, called self-excited oscillations and perceived as frictional noise or squeal.

Different types of these vibrations have been studied experimentally by several investigators, including Soom and Kim [1983], Ko and Brockley [1967], Bhushan [1980], and Arenov, D'Souza, Kalpakjann, and Shareef [1985]. The first numerical study of this phenomenon was presented by Oden and Martins [1985], who estimated frictional stability of a rigid block sliding on a rigid surface by the analysis of eigenvalues of linearized equations of motion of the system (equation 5.10 in this paper). This line of research was continued by Becker and Tworzydło [1988], who have identified the mechanism of instability as the *coupling between normal and rotational vibrations* of the slider in presence of friction.

The detailed analysis of the problem of dynamic instability and self-excited oscillations will be presented in Part II of this work.

No	Phenomenon	Description	Group	Reason - model	Result	When active	Importance
1	Dynamic instability of the system	Oscillations grow until normal jumps occur	B	coupling between normal and rotational oscillations	-self excited oscillations -stick-slip II kind	When $\omega_0 = \omega_y$ and sliding occurs	++
2a	Rate dependence of static friction	μ is smaller for higher rates of application of tangential force	A/B	plastic increase of approach and contact area ; initial microdispl.	stick-slip I kind	beginning of slip in absence of normal oscillations	++
2b	Time dependence of static friction	μ increases with time of static contact	A	creep increase of the contact area; diffusion	stick-slip I kind	static contact - no slip or normal oscillations	-
3	Dissipation during normal oscillations	energy is dissipated during oscillations due to viscoplastic deformations	A/B	a) plastic dissipation b) viscoelastic and viscoplastic damping	-damping of normal osc. -decrease of coef. of fric.	whenever normal oscillations occur	++
4	Normal jumps of a slider	two surfaces loose contact	B	-growth of oscillat. in unstable cases -interlocking of asper. at begin. of slip	-decrease of coef. of fric. -excitation of oscillations	-instable cases -beginning of slip after stick	++
5	Load dependence of the coefficient of frict.	μ changes with the normal load	A	-breaking through oxide layer -plastic failure at very high loads	changes of apparent μ when oscillations occur	whenever normal oscillations occur	+
6	Dynamic interlocking of imperfections	Imperfections of both surfaces impinge on each other	A/B	roughness of a surface	-excitation of normal oscillations - lift force	when macroscopic sliding occurs	++

Table 1: Summary of the most important phenomena of dynamic friction.

Here we will only mention that once the self-excited oscillations have occurred, they can be the reason of other occurrences, like the reduction of the coefficient of friction, stick-slip motion or normal jumps of the slider.

2.2 Rate and Time Dependence of the Static Coefficient of Friction

The time or rate dependence of the static coefficient of friction are typical phenomena of group A. Historically, the first one observed by researchers was time dependence of the static friction. It was believed (see, e.g., Sampson, et al. [1943], Rabinowicz [1958], Brockley and Davis [1968]) that at zero time of stick the static coefficient of friction was equivalent to the kinetic coefficient, but then its value increased with the time of stick.

This growth of the coefficient of friction was explained by increase of the real contact area due to viscoplastic deformations of asperities under the static normal load. Various formulas expressing the time change of μ_s were proposed. The one in best agreement with the experimental observations was proposed by Kato, Sato, and Matsubayachi [1972] in the form:

$$\mu_s(t_c) = \mu_{s\infty} - (\mu_{s\infty} - \mu_0)e^{-\gamma(t_c)^m} \quad (2.1)$$

where:

t_c - time of static contact,

μ_s - the static coefficient of friction,

μ_0 - the static coefficient of friction at zero time of stick,

$\mu_{s\infty}$ - the asymptotic value of the static coefficient of friction after long time of static contact,

γ, m - parameters.

However, more recent experimental works of Johannes, Green, and Brockley [1973] and Richardson, Nolle [1976] showed that the phenomena assigned to the time dependence of the static friction should rather be interpreted as the rate dependence of the static coefficient of friction. From specially designed experiments these authors derived a conclusion that the final value of the static coefficient of friction depends primarily on the rate of application of the tangential load, i.e., the higher the rate, the smaller the coefficient of friction.

The typical μ_s versus $\dot{\Phi}$ curve obtained by Richardson, Nolle [1976] is presented in Figure 1. This curve can be described by equation (2.1), provided that the rate of application of the tangential load is constant and that there are no interruptions to this loading. For the more general cases of loading the authors suggested reformulation of equation (2.1) in terms

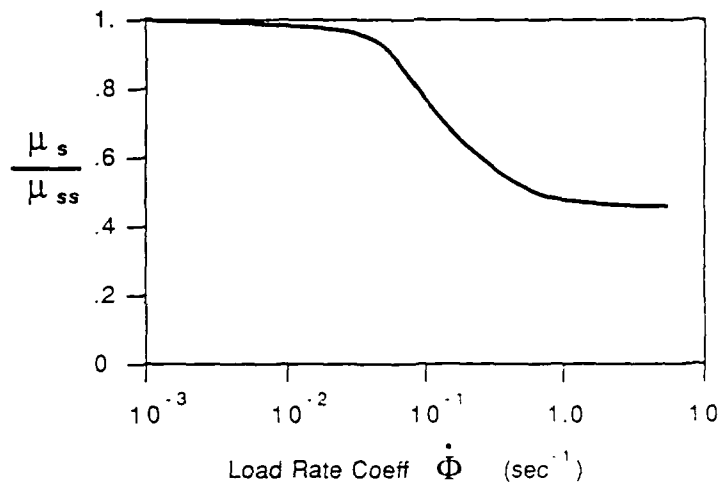


Figure 1: The rate dependence of the static coefficient of friction — Richardson, Nolle [1976]

of loading rate as the independent variable (they did not propose the form of this modified equation). In the experimental work which they presented, the authors did not suggest any explanation of the mechanism of the observed phenomenon, which is of a somewhat different nature than that of time dependence of static friction. It seems that the most natural explanation is the viscoplastic growth of the area of junctions due to combined tangential and normal loads, described within the purely plastic range by the early theory of junction growth of Bowden and Tabor [1964]. Since the process is of viscoplastic nature, one can expect smaller junction growth, i.e., the smaller coefficient of friction, at high rates of the tangential load.

In the context of kinetic friction, which is of primary interest in this work, rate dependence of the static coefficient of friction can be important in the case of stick-slip motion. In this case rate dependence can strongly affect the value of tangential force, at which the slip occurs. This, in turn, can influence the nature of motion of the slider, the value of apparent kinetic coefficient of friction, etc.

2.3 Dissipation of Energy During Normal Oscillations

The normal oscillations of the slider are usually accompanied by severe deformations of the surface asperities, during which some dissipation of energy occurs.

Generally this dissipation is due to:

1. visco-elasto-plastic deformations of asperities and wear debris, and

2. viscous properties of lubricants.

The main dynamic effect produced by this dissipation is the damping of normal oscillations of the slider. Moreover, since most of the dissipation occurs when two bodies approach each other, the additional effect of oscillations is a slight rising of the slider, which in turn results in certain reduction of the real contact area and of the coefficient of friction. This effect was observed experimentally by Tolstoi [1966].

It should be pointed out that the viscoplastic deformation of asperities is of a complex nature and therefore difficult to model exactly in the phenomenological constitutive model of the interface. However, for the sake of general modeling of vibrations of the slider the damping effect produced by this deformation can be represented by a nonlinear damping term in the constitutive law of the interface (see Hunt and Crossley [1975], Oden and Martins [1985], and Tworzydło and Becker [1988]). The corresponding model will be presented in section 4.3 of this paper.

It should be pointed out that modeling of dissipation of energy on the interface by the damping term only is a very rough approximation of real phenomena. Moreover, due to oversimplicity of this model and lack of reliable experimental data, estimating values of coefficients of damping is very difficult. Results of impact experiments (Hunt, Crossley [1975]) are, due to completely different range of normal velocities, inapplicable in the case of vibrations accompanying frictional sliding.

2.4 Normal Jumps of a Slider

Normal jumps of a slider are a very specific feature of the dynamic friction and need some special attention. They can occur in two typical situations:

- in the case of high-amplitude self-excited oscillations,
- at the very beginning of the sliding after the static contact of two surfaces (slip after stick).

The first type of jumps, which are easily observed in many experiments on frictional vibrations (see, e.g., Ko and Brockley [1970], Kato, et al. [1970], and Aronov, et al. [1983, 1984]) are caused by the growth of amplitudes of normal oscillations in the dynamically unstable cases.

The second kind of jumps was observed in the experiments of Grigorova, Tolstoi [1966], Tolstoi [1967], and Tolstoi, Borisova, and Grigorova [1971], who investigated stick-slip motion with precise analysis of normal displacements. They observed that *"forward movements of the slider invariably occurred in strict synchronism with its upward jumps."*

These jumps (see Fig. 2) at the very beginning of the sliding phase are significantly different from the normal oscillations occurring during steady-sliding, which usually do not

cause loss of contact between the two surfaces. The importance of these jumps results from the fact that they provide strong initial impulse for the normal oscillations of the slider and thereby activate most of the other features of the dynamic friction. Therefore it would be of importance — especially in the analysis of stick-slip motion — to model these jumps in the numerical analysis, for example in the form of additional perturbation of normal displacements. Some examples of this approach were presented by Becker and Tworzydło [1988].

2.5 Load Dependence of the Coefficient of Friction

Load dependence of the static coefficient of friction, apart from its significance in the purely static friction, also produces some additional effects in the case of dynamic friction or — more generally — whenever normal oscillations occur. As it was shown by Becker and Tworzydło [1988], this effect is as follows: under constant normal load the load dependence of μ_s is the reason of change of apparent kinetic coefficient of friction μ_k in presence of normal oscillations. In particular, if μ_s decreases with normal load, then in presence of normal oscillations the apparent coefficient of friction decreases (compared with the static friction). Reversely, if μ_s increases with the normal load, the macroscopic coefficient of friction increases in presence of normal oscillations (in the absence of other effects).

It is of importance to note that at the current state of knowledge taking into account the load dependence of the coefficient of friction in practical applications is somewhat difficult. The reason is that load dependence of μ_s has a different nature for different materials and types of surfaces. Even for the same material the data obtained by different researchers can differ significantly and there is no agreement even with regard to the general nature of load dependence. For example, for steel or iron surfaces some researchers (Wilson [1952], Bowden and Taylor [1964] and other sources referred in this book) observed a slow decrease in the coefficient of friction at light loads, rapid fall at moderate loads and then almost constant value (Fig. 3). On the contrary, other authors suggest slight increase of the coefficient of friction at light and moderate loads and drop at high loads (see, e.g., Tolstói [1967], Nolle and Richardson [1974], Buckley [1977], Bay and Wendheim [1976], Broniec and Lenkiewicz [1980]). Apart from this difficulty, in the computational part of this work we will qualitatively illustrate the influence of the load dependence of friction on the value of the dynamic coefficient of friction (for the particular setup considered here).

2.6 Dynamic Interlocking of Imperfections

The dynamic interlocking of imperfections is a key factor in frictional sliding. It is so not only because it strongly affects the frictional resistance of the interface, but also because it provides permanent excitation of normal vibrations of the slider. It is this second role that

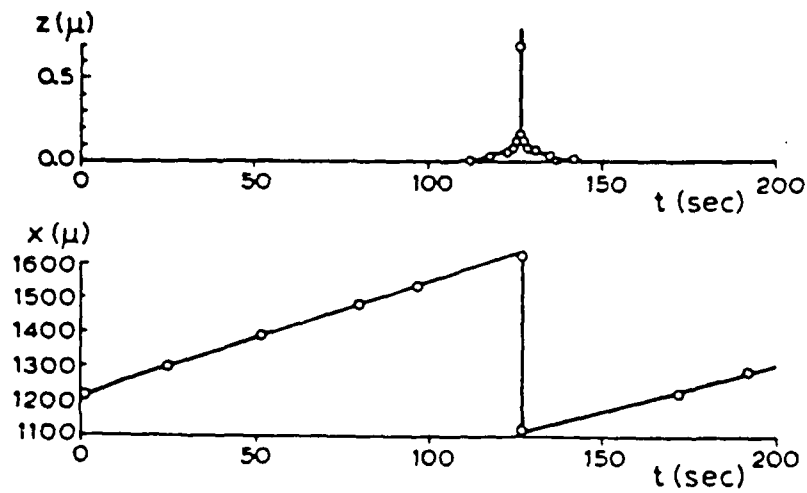


Figure 2: Normal jumps of the slider in the stick-slip motion — Tolstoi [1967]

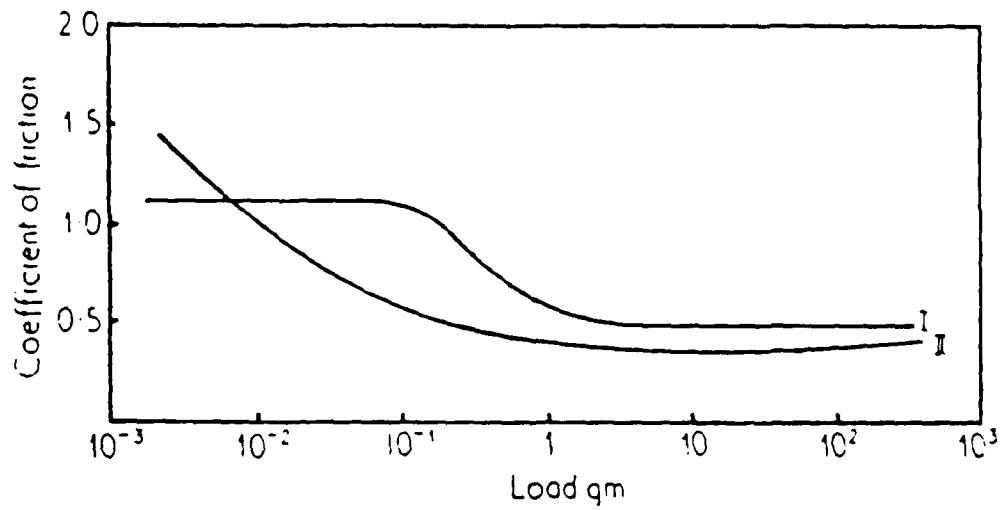


Figure 3: Load dependence of the static coefficient of friction of steel — Wilson [1952]

we would like to emphasize in this paragraph.

The first researcher to clearly show the existence and importance of the normal motion of a slider was Tolstoi [1966]. Then this phenomenon was studied experimentally by Antoniou, et al. [1976], Sakamoto, et al. [1980], Soom and Kim [1983 a, b] and Aronov, et al. [1984]. The explanation of the source of these vibrations by interlocking of imperfections of the surface and propositions of modeling of this interlocking were presented by Anand and Soom [1984], Lanchon, et al. [1985] and Soom and Chen [1986]. Using sinusoidal and quasi-random model of waviness of the surface and simple model of the slider they simulated dynamic effects observed in experiments.

It is of importance to observe that in the case of self-excited oscillations the interlocking of imperfections provides initial impulse for oscillations, but then vibrations reach much higher levels than those forced by waviness of the surface only (like in so-called steady-sliding). Therefore, in this case, which is of primary interest in this paper, there is no need for precise modeling of excitation due to imperfections of the surface and it suffices to model only the initial impulse, which triggers further self-excited vibrations of the slider.

3 General Presentation and Formulation of the Boundary-Value Problem

The primary object of our interest in this report is a typical pin-on-disk apparatus, which is a good representative of a broad class of mechanical sliding systems. The computational model of such a setup should, in general, be three-dimensional.

However, if one is interested primarily in the motion of the pin and the deformation of the disk in the vicinity of pin, the behavior of the system can be modeled by a two-dimensional model of an elastic body A in contact with another elastic body B, moving with a prescribed velocity \dot{U} (Fig. 4).

It is of importance to note that we are only interested in deformation of certain part of body B, which is currently in the neighborhood of contact zone. If one focuses attention on this portion, it can be represented by a hypothetical frame F, with material matter "flowing" through the frame with the prescribed velocity. The exact definition of the frame F and reformulation of equations of body B in terms of deformation of frame will be presented in the next section. Here we will present the most general assumptions for the model presented in Fig. 4, which are as follows:

- deformations of bodies A and B are infinitesimal,
- materials of the two bodies are linearly elastic,
- initial gap between A and B is small.

A more detailed specification of assumptions will be presented in the subsequent paragraphs.

In the formulation of a boundary-value problem for the model of Fig. 4 we will assume that all vectors and tensor fields are well defined and sufficiently smooth, so that all necessary derivatives are well defined. Our notation corresponds to the general notation of continuum mechanics. In particular, boldface symbols are reserved for points, vectors or tensors, and symbols with indices represents their components in certain coordinate systems. When applicable, capital letters correspond in general to reference configuration while small letters correspond to (one of) current configurations. Moreover, whenever it is essential, the configuration will be indicated by explicitly stating the time variable.

Since in our analysis we deal with two bodies and a specially defined frame, objects assigned to particular bodies will be indicated by a superscript A, B, or F unless this assignment is clear from the context.

Detailed explanation of all symbols will consequently be presented in the text.

3.1 Equations for Body A

The body A, which represents the slider, is assumed to be linearly elastic and subject to small deformations. Therefore the equations of motion for body A are of the standard form:

$$\operatorname{div} \mathbf{T} + \mathbf{b} = \rho \mathbf{a} \quad \text{in } \Omega^A \times [0, T] \quad (3.1)$$

where ρ is the density of material, \mathbf{b} is the body force vector, \mathbf{a} is the acceleration vector $\mathbf{a} = \ddot{\mathbf{u}}$ and \mathbf{T} is the Cauchy stress tensor, calculated in terms of displacements as:

$$\mathbf{T} = \mathbf{E}(\nabla \mathbf{u})$$

where \mathbf{E} is the fourth rank elasticity tensor, satisfying standard conditions of symmetry, boundedness and ellipticity.

In addition to the above equation one can define a variety of boundary conditions, the most typical being proscribed displacements \mathbf{u}_D , prescribed tractions \mathbf{f} or elastic support with proscribed displacements \mathbf{u}_E :

$$\begin{aligned} \mathbf{u} &= \mathbf{u}_D & \text{on } \Gamma_D^A \\ \mathbf{T}\mathbf{n} &= \mathbf{f} & \text{on } \Gamma_F^A \\ \mathbf{T}\mathbf{n} &= \mathbf{K}_E(\mathbf{u} - \mathbf{u}_E) & \text{on } \Gamma_E^A \end{aligned} \quad (3.2)$$

Moreover, we need to specify initial conditions:

$$\begin{aligned} \mathbf{u}(0) &= \mathbf{u}_0 & \text{in } \Omega^A \\ \dot{\mathbf{u}}(0) &= \mathbf{v}_0 & \text{in } \Omega^A \end{aligned} \quad (3.3)$$

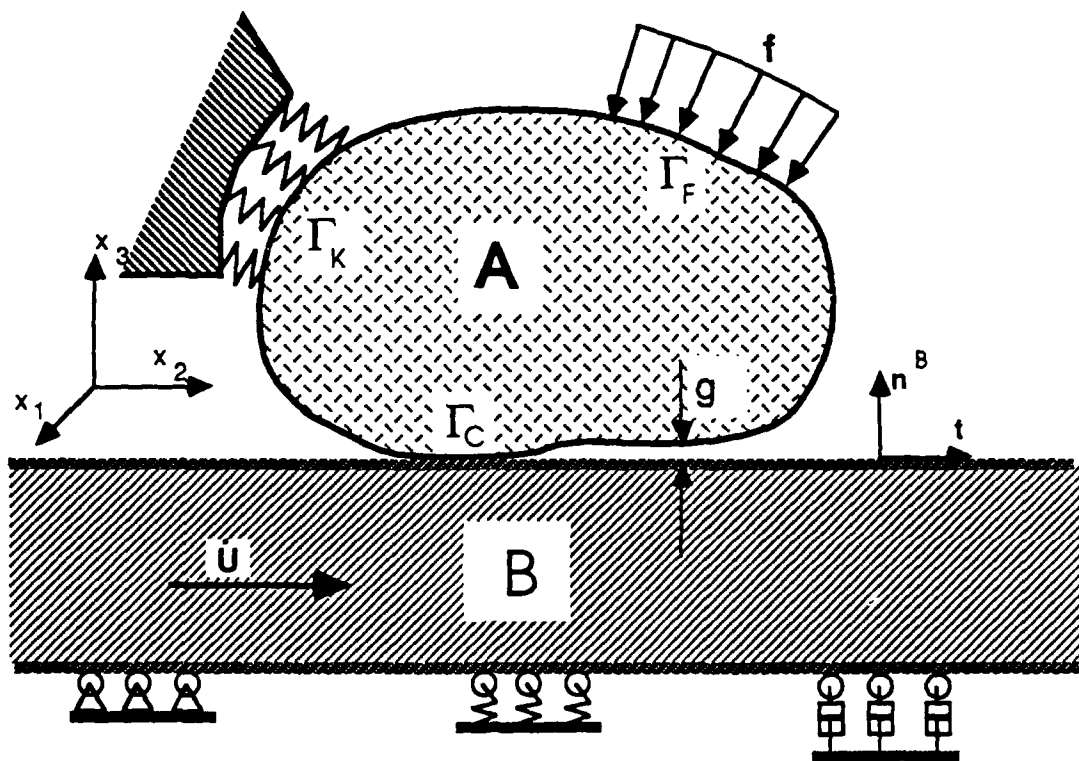


Figure 4: Two elastic bodies in sliding contact

It should be noted that, since rigid translation does not change the deformation gradient, the above equations are also valid in the case of finite translations superimposed on infinitesimal deformations of body A.

3.2 Representation of the Moving Body B by a Frame F

In this section we formulate equations of body B in terms of frame F. We assume the following:

1. Strains and rotations in body B are infinitesimal.
2. The body is subject to the overall rigid translation with prescribed, constant velocity \dot{U} . The current deformation of body B at time t is a composition of this overall motion and infinitesimal, dynamic deformations due to contact with body A.
3. Properties of the body B are homogenous along the direction of \dot{U} .
4. Body B extends infinitely along \dot{U} and its boundaries in the reference configuration are parallel to \dot{U} .

The general idea of introducing the frame F is to extract the above-mentioned rigid motion from the deformation of body B, so that the frame would stay in contact with body A and represent the essential deformation of body B due to dynamic contact forces.

To make this idea more precise, consider the system in the reference configuration at time $t_0 = 0$ and in deformed configuration at time t (see Fig. 5).

In particular, consider the segment of the body $\hat{B}(t_0)$, which at the reference time $t_0 = 0$ is "under" body A. Let us identify frame F (t_0) with this segment so that

$$\mathbf{x}^F(\mathbf{X}^F, t_0) = \mathbf{X}^F(t_0) = \mathbf{X}^{\hat{B}}(t_0) = \mathbf{x}^B(\mathbf{X}^{\hat{B}}, t_0) \quad (3.4)$$

where $\mathbf{x}(\mathbf{X}, t)$ is the current position of a particle \mathbf{X} and superscripts B and F refer to body B and frame F, respectively.

At time $t > t_0$ the portion \hat{B} moves away from the neighborhood of A, but another segment comes into contact with A. In order to have a reference frame still in contact with A, we require that at time t the frame corresponds to this new segment, denoted by B .

This means that we want current position of frame F to correspond to the configuration of B,

$$\mathbf{x}^F(\mathbf{X}^F, t) = \mathbf{x}^B(\mathbf{X}^B, t) \quad (3.5)$$

in particular of the material part which is currently in contact with body A. This is guaranteed by selecting:

$$\begin{aligned} \mathbf{X}^B &= \mathbf{X}^F - \dot{\mathbf{U}}t & (a) \\ \text{or} & & (3.6) \\ \mathbf{X}^F &= \mathbf{X}^B + \dot{\mathbf{U}}t & (b) \end{aligned}$$

Equations (3.5) and (3.6) represent basic kinematic relation between frame F and body B. The result of such a definition is that the frame F is not an Eulerian frame fixed in space, but represents the essential part of the motion of body B (like its vibrations), with only the overall prescribed motion extracted from the velocity field.

The above concept of a frame resembles the so-called Arbitrary Lagrangian-Eulerian (ALE) description of deformations, developed for the fluid-structure interaction by Kennedy and Belytschko [1981], Donea, Guiliani and Holleux [1982] or Schreurs, Veldpous and Brekelmans [1985]. The basic difference is in our approach to the relation between the two control systems, given in the *references*, rather than in the current configuration.

3.2.1 Kinematics of a Frame

Relations (3.5) and (3.6) are the basis for the kinematic description of body B in terms of frame F, in particular relating displacements, velocities and accelerations in the two descriptions.

Displacements of body B and frame F are defined by

$$\begin{aligned} \mathbf{u}^F &= \mathbf{x}^F - \mathbf{X}^F \\ \mathbf{u}^B &= \mathbf{x}^B - \mathbf{X}^B \end{aligned} \quad (3.7)$$

which, after application of (3.5) and (3.6) leads to the correspondence

$$\begin{aligned} \mathbf{u}^B(\mathbf{X}^B, t) &= \mathbf{u}^F(\mathbf{X}^F, t) + \dot{\mathbf{U}}t \\ \text{or} & \\ \mathbf{u}^B(\mathbf{x}, t) &= \mathbf{u}^F(\mathbf{x}, t) + \dot{\mathbf{U}}t \end{aligned} \quad (3.8)$$

It is easy to observe from (3.6) that

$$\nabla_B(\cdot) \approx \nabla_F(\cdot) \quad (3.9)$$

where ∇_B and ∇_F represent material gradients w/r to particles of a body and frame, respectively. Then from equation (3.8) one can conclude that

$$\nabla_B \mathbf{u}^B = \nabla_F \mathbf{u}^F \quad (3.10)$$

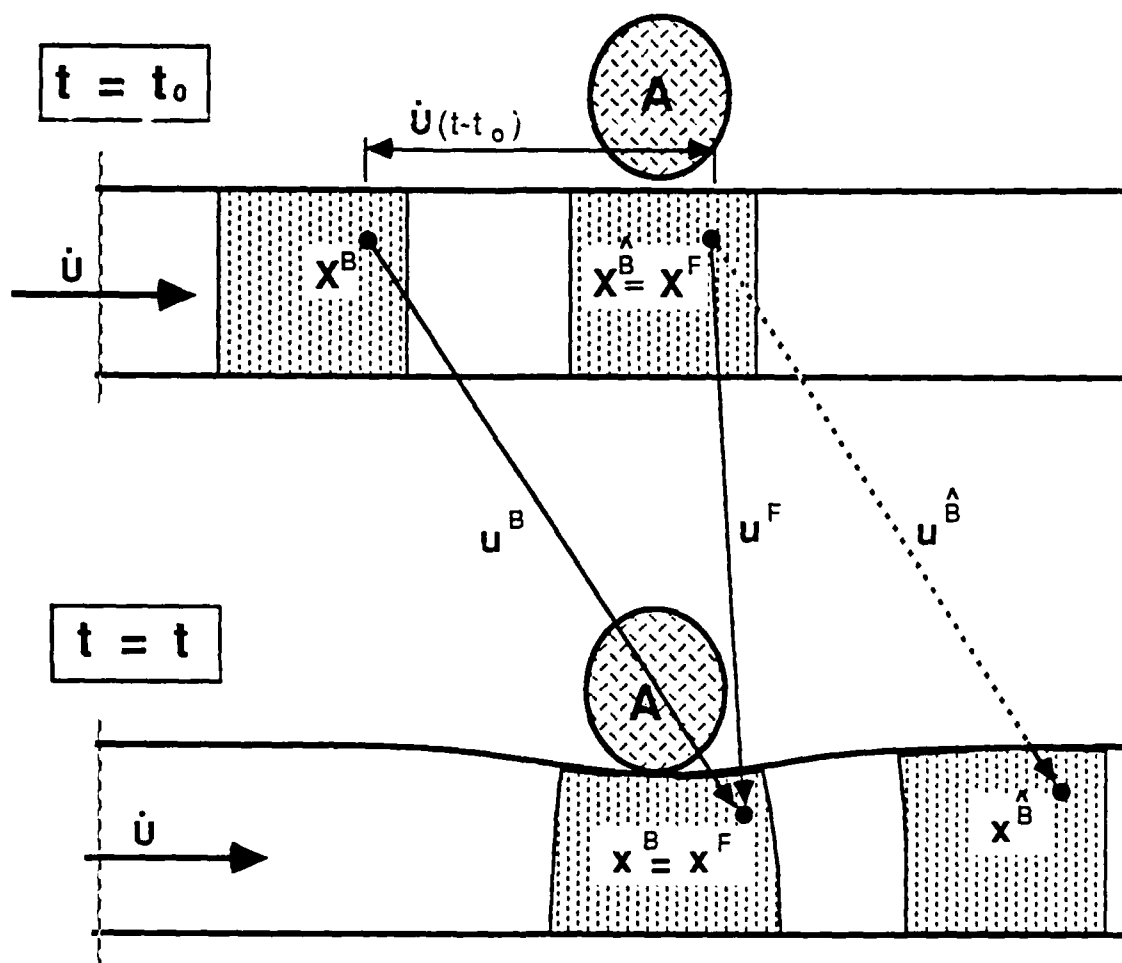


Figure 5: Motion of body B and frame F

This means that the strains calculated for the frame are identical as the strains in the body (the part that currently corresponds to the frame).

Velocities and accelerations of particles of body B are expressed in terms of appropriate time derivatives of vector fields defined on B, for example a vector field $\mathbf{w} = \mathbf{w}(\mathbf{x}^B, t)$ the time derivative in

$$\left. \frac{D}{Dt} \right|_{\mathbf{X}^B \text{ fixed}} \mathbf{w}(\mathbf{X}^B, t) = \frac{\partial}{\partial t}(\mathbf{X}^B, t)$$

In order to express this material time derivative for fields defined in frame F one can use chain rule combined with the relation (3.6) to obtain:

$$\frac{D}{Dt} \mathbf{w}(\mathbf{X}^F, t) = \frac{\partial}{\partial t} \mathbf{w}(\mathbf{X}^F, t) + (\nabla_F \mathbf{w}) \dot{\mathbf{U}} \quad (3.11)$$

or for scalar fields

$$\frac{D}{Dt} \alpha(\mathbf{X}^F, t) = \frac{\partial}{\partial t} \alpha(\mathbf{X}^F, t) + (\nabla_F \alpha) \cdot \dot{\mathbf{U}} \quad (3.12)$$

Application of formula (3.11) to positions \mathbf{x}^B of body B and noting that $\mathbf{x}^B(\mathbf{X}^B, t) = \mathbf{x}^F(\mathbf{X}^F, t)$ leads to the following expression of velocities of body B in terms of velocities of "particles" of a frame F:

$$\mathbf{v}^B = \mathbf{v}^F + (\nabla_F \mathbf{x}^F) \dot{\mathbf{U}} \quad (3.13)$$

If we want to express the above formula in terms of displacements rather than positions of "particles" of a frame, then with the use of (3.7 a) we obtain the relation

$$\mathbf{v}^B = \mathbf{v}^F + \dot{\mathbf{U}} + (\nabla_F \mathbf{u}^F) \dot{\mathbf{U}} \quad (3.14)$$

In an analogous way one can express accelerations of particles of body B, defined by

$$\mathbf{a}^B = \left. \frac{D}{Dt} \right|_{\mathbf{X}^B \text{ fixed}} \mathbf{v}^B$$

in terms of accelerations of "particles" of a frame, defined as

$$\mathbf{a}^F = \left. \frac{D}{Dt} \right|_{\mathbf{X}^F \text{ fixed}} \mathbf{v}^F = \frac{\partial}{\partial t} \mathbf{v}^F(\mathbf{X}^F, t)$$

Substitution of formula (3.14) and application of (3.11) yields the final expression:

$$\mathbf{a}^B = \mathbf{a}^F + 2(\nabla_F \mathbf{v}^F) \dot{\mathbf{U}} + \nabla_F ((\nabla_F \mathbf{u}^F) \dot{\mathbf{U}}) \dot{\mathbf{U}} \quad (3.15)$$

This formula shows that accelerations of material particles depend on both the velocity field and deformation (curvatures) of frame F.

3.2.2 Conservation Laws in Frame Description

In this paragraph we will derive basic conservation laws for body B, expressed in terms of the reference frame F. In our derivations we will consequently use the assumption that the deformation of body B consists of infinitesimal deformation superimposed on rigid translation.

Prior to evaluation of conservation laws let us observe that deformation gradients for both body B and frame F can be presented as:

$$\mathbf{F}^B = \mathbf{I} + \nabla_B \mathbf{u}^B, \quad \mathbf{F}^F = \mathbf{I} + \nabla_F \mathbf{u}^F$$

Recalling observation (3.10) from the previous section, we conclude that the deformation gradients for body B and frame F are equal

$$\mathbf{F}^B = \mathbf{F}^F \quad (3.16)$$

and for small gradients of displacements they are arbitrarily close to the identity tensor. This also means that

$$\mathbf{F}^B = \mathbf{F}^F = \mathbf{I} + o(\nabla_F \mathbf{u}^F) \quad (3.17)$$

$$\det \mathbf{F}^B = \det \mathbf{F}^F \cong 1 \quad (3.18)$$

In view of the above result, conservation of mass for body B takes the form:

$$\rho(\mathbf{X}^B, 0) = \rho(\mathbf{x}^B, t) \quad (3.19)$$

where $\mathbf{x}^B = \mathbf{x}^B(\mathbf{X}^B, t)$.

It is easy to observe that from (3.5) we have $\rho(\mathbf{x}^B, t) = \rho(\mathbf{x}^F, t)$. Moreover, from (3.6) one can easily deduct that for homogenous distribution of density along $\dot{\mathbf{U}}$, $\rho(\mathbf{x}^B, 0) = \rho(\mathbf{x}^F, 0)$.

Then the final form of conservation of mass in the frame description is:

$$\rho(\mathbf{X}^F, 0) = \rho(\mathbf{x}^F, t) \quad (3.20)$$

The equation of momentum for body B is of the form:

$$\operatorname{div}_B \mathbf{T}^B + \mathbf{b}^B = \rho^B \mathbf{a}^B$$

where, in view of (3.17) and (3.18), \mathbf{T}^B is a Cauchy stress tensor, \mathbf{b}^B is the density of body forces in the current configuration, and \mathbf{a}^B is the acceleration of particles of body B in the current configuration.

If we observe that:

- current configurations of body B and frame F are identical,

- corresponding deformation gradients are identical (3.16),
- physical properties for corresponding particles of Body and Frame are the same (by the assumption of homogeneity along \dot{U} ,

and use (3.15), then we arrive at the equations of motion in the frame description

$$\operatorname{div}_F \mathbf{T}^F + \mathbf{b}^F = \rho^F [\mathbf{a}^F + 2(\nabla_F \mathbf{v}^F) \dot{U} + \nabla_F ((\nabla_F \mathbf{u}^F) \dot{U}) \dot{U}] \quad (3.21)$$

3.2.3 Boundary Conditions for Body B

Boundary conditions for body B depend on the actual system to be modeled. However, due to specific motion of this body, one can formulate a general type of boundary conditions applicable in this case.

In general, since body B is constantly sliding, its boundaries parallel to the direction of motion can be supported by a variety of rollers or bearings; rigid, elastic or with damping properties. The corresponding equations expressed in terms of deformation of frame are of the form (compare with Fig. 4)

$$\left\{ \begin{array}{ll} \mathbf{u}^F \cdot \mathbf{n} = 0 & \text{on } \Gamma_N^F \times [0, T] \\ (\mathbf{T}^F \mathbf{n}) \cdot \mathbf{n} = -k \mathbf{u}^F \cdot \mathbf{n} & \text{on } \Gamma_E^F \times [0, T] \\ (\mathbf{T}^F \mathbf{n}) \cdot \mathbf{n} = -d \mathbf{v}^F \cdot \mathbf{n} & \text{on } \Gamma_D^F \times [0, T] \end{array} \right. \quad (3.22)$$

It was generally assumed that body B extends infinitely along direction of \dot{U} . However, we would usually model only the segment of body B in the vicinity of the contact zone. Then the simplest boundary conditions for boundaries perpendicular to \dot{U} would be

$$\mathbf{u}^F = \mathbf{0} \quad \text{on } \Gamma_I^F \times [0, T]$$

which means that particles of B have velocity \dot{U} on these boundaries.

3.3 Model of the Interface

3.3.1 Definition and Assumptions

The interface between the two bodies A and B is — in our approach — the hypothetical medium of very small thickness, covering surfaces of contacting bodies. It represents properties of asperities, oxide and contaminant layers on the surface of real bodies. The boundaries of the surface layers may, for example, correspond to surfaces passing through the tops of higher asperities and bottoms of the deepest grooves on the surface (see Fig. 6).

The most important assumptions for the contact surfaces and interface are the following:

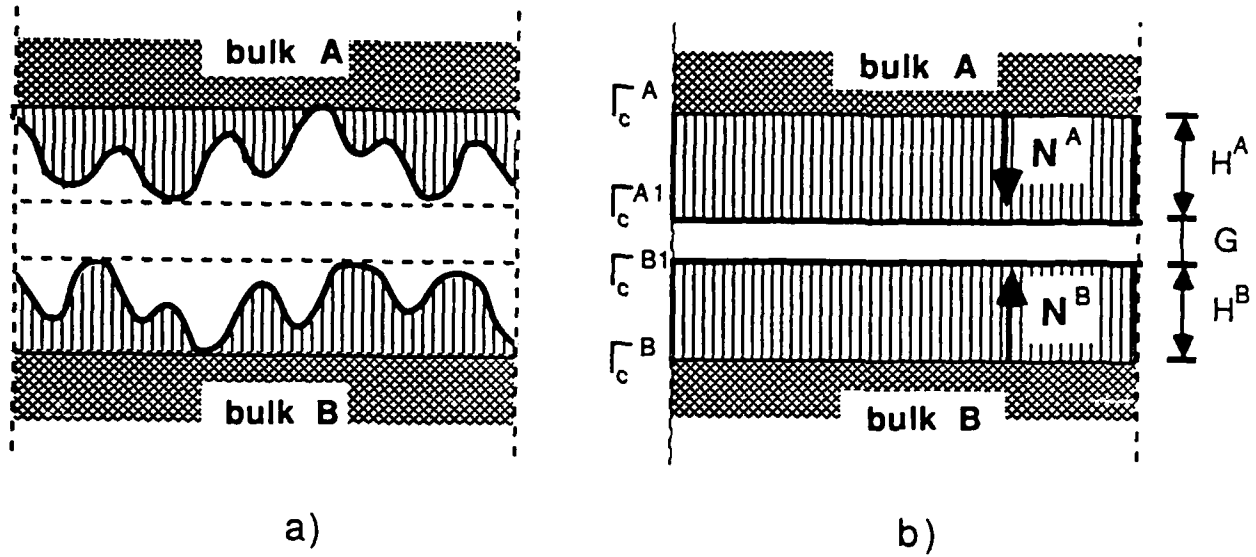


Figure 6: Continuum model of the frictional interface

1. All surfaces $(\Gamma_c^A, \Gamma_c^{A1}, \Gamma_c^B, \Gamma_c^{B1})$ are smooth (C^1)
2. Thickness of both layers is small compared to the dimensions of the body

$$1 + \max \left(\frac{H^\Delta}{\rho^\Delta}, \frac{H^\Delta}{L^\Delta} \dots \right) \approx 1 \quad \text{for } \Delta = A, B$$

where ρ^Δ is the local radius of curvature and L^Δ is a typical size of the body

3. Changes of thickness are small

$$1 + \frac{\partial H^\Delta}{\partial S^\Delta} \approx 1 \quad \text{for } \Delta = A, B$$

where S parameters the contact surface.

Moreover, we will assume in this work that the initial gap between the surfaces is small and known *a priori* (for details of calculation of gap one can refer for examples to Kikuchi, Oden [1988] and references therein).

After applications of loads on the system the two surfaces come into contact, so that sections of $\Gamma_c^{A1}, \Gamma_c^{B1}$ become a common contact boundary Γ_c .

The important conclusion of all the above assumptions and the assumption of small rotations of bodies A and B is that in the contact zone the normal vectors to both sides can be considered to be identical:

$$\mathbf{N}^A(\mathbf{X}^A, 0) = \mathbf{n}^A(\mathbf{X}^A, t) = -\mathbf{N}^B(\mathbf{X}^B, 0) = -\mathbf{n}^B(\mathbf{X}^B, t) \quad (3.23)$$

where particle \mathbf{X}^B is located opposite particle \mathbf{X}^A in the current configuration.

Moreover, from the macroscopic point of view we can identify the surfaces Γ_C^A , Γ_C^B and Γ_C in the sense that various objects can be referred to any of these surfaces and that the surface integration can be carried out on any of them.

3.3.2 Kinematics of the Interface

If one takes into account the assumptions presented in the previous paragraph, then approach (or penetration) on the interface can be calculated as:

$$a(\mathbf{X}^A, t) = u_N^R(\mathbf{X}^A, t) - G(\mathbf{X}^A) \quad (3.24)$$

where U_N^R is the relative normal displacement of body B with respect to body A:

$$u_N^R(\mathbf{X}^A, t) = \mathbf{u}^R(\mathbf{X}^A, t) \cdot \mathbf{n}^A$$

with

$$\mathbf{u}^R(\mathbf{X}^A, t) = \mathbf{u}^B(\mathbf{X}^B, t) - \mathbf{u}^A(\mathbf{X}^A, t)$$

where \mathbf{X}^B is the particle opposite \mathbf{X}^A on Γ_C . The sliding vector is defined as the projection of the relative displacement on the surface parallel to Γ_C :

$$\mathbf{d}(\mathbf{X}^A, t) = \mathbf{u}^R(\mathbf{X}^A, t) - u_N^R(\mathbf{X}^A, t)\mathbf{n}^A$$

and in our two-dimensional case represented by the scalar

$$d = u_T^R = \mathbf{u}^R(\mathbf{X}^A, t) \cdot \mathbf{t} \quad (3.25)$$

where \mathbf{t} is the vector parallel to Γ_C . In our description, which represents body B by the frame F, we should use formulas (3.8) to express penetration and sliding distance in terms of displacements of a frame.

Application of (3.8) and observing that by (3.23) \mathbf{n}^A is orthogonal to $\dot{\mathbf{U}}$, renders the following expressions:

$$a(\mathbf{X}^A, t) = (\mathbf{u}^F(\mathbf{X}^F, t) - \mathbf{u}^A(\mathbf{X}^A, t)) \cdot \mathbf{n}^A - G(\mathbf{X}^A) \quad (3.26)$$

$$d(\mathbf{X}^A, t) = (\mathbf{u}^F(\mathbf{X}^F, t) - \mathbf{u}^A(\mathbf{X}^A, t)) \cdot \mathbf{t} + \dot{\mathbf{U}}t \quad (3.27)$$

In the constitutive equations of the interface we will also use rate of approach \dot{a} and sliding velocity \dot{d} , which are calculated as appropriate time derivatives. Before carrying out these calculations it is important to observe that for fixed \mathbf{X}^A location of \mathbf{X}^F which is opposite to \mathbf{X}^A changes.

If we recall our basic assumptions (small strains and rotations, small initial gap, orthogonality of \mathbf{n} to $\dot{\mathbf{U}}$ on Γ_C) and locate \mathbf{X}^F by means of normal projection of \mathbf{X}^A on Γ_C , then:

$$\frac{d\mathbf{X}^F(\mathbf{X}^A, t)}{dt} = (v_T^A - v_T^F)\mathbf{t}$$

where v_T^A, v_T^F are tangential velocities of particles $\mathbf{X}^A, \mathbf{X}^F$. Then the time derivative of \mathbf{u}^F in (3.27) and (3.27) can be expressed as

$$\frac{d\mathbf{u}^F(\mathbf{X}^F, t)}{dt} = \frac{\partial \mathbf{u}^F(\mathbf{X}^F, t)}{\partial t} + (\nabla_F \mathbf{u}^F) \frac{d\mathbf{X}^F}{dt}$$

where $\partial \mathbf{u}^F / \partial t = \mathbf{v}^F$.

Using the above formulas we obtain the following expressions for the rate of approach and relative sliding velocity:

$$\dot{a} = (v_N^F - v_N^A) + (v_T^F - v_T^A) (\nabla_F \mathbf{u}^F) \mathbf{t} \cdot \mathbf{n}^A \quad (3.28)$$

$$\dot{d} = (v_T^F - v_T^A) (1 + (\nabla_F \mathbf{u}^F) \mathbf{t} \cdot \mathbf{t}) + |\dot{\mathbf{U}}| \quad (3.29)$$

Moreover, if we recall in (3.29) that gradient $\nabla_F \mathbf{u}^F$ is small (in the sense of norm of linear operators) then we obtain simplified expression for \dot{d}

$$\dot{d} = (v_T^F - v_T^A) + |\dot{\mathbf{U}}| \quad (3.30)$$

We cannot, however, make this simplification in (3.28) because we do not know *a priori* the proportion between normal and tangential velocities of body A and frame F.

3.3.3 Constitutive Relations for the Interface

When formulating constitutive equations for the interface one should observe that the interface consists of two separate layers, often of different properties. Therefore, in the general approach, constitutive properties should be estimated for each layer separately, and the overall equations of the interface should be based on these properties.

On the other hand, we are usually interested in properties of interface as a whole, and moreover, such is the nature of experimental results. Therefore one can define material constants which represent the interface as a whole, remembering that for dissimilar materials one should define these constants for each pair of materials separately.

The constitutive equations of the interface applied in this work are based on the Oden/Martins model and consist of two basic equations:

- normal interface law,

- friction law.

The normal response law is of the form

$$-\sigma_N = -c_N a^{m_N} + b_N a^{l_N} \dot{a} \quad (3.31)$$

which corresponds satisfactorily to experimental observations at moderate loads, reviewed by Back, Burdekin and Cowley [1973].

The first term in this equation represents elastic response of the interface, while the second damping term models dissipation of energy during normal oscillations. As mentioned earlier, the nature of this dissipation is fairly complex and the above term is only supposed to model the overall damping effect produced by the viscous properties of surface asperities, contaminants and lubricants. The coefficients b_N, l_N are usually assumed to be constant — see Hunt and Crossley [1975] and Oden and Martins [1986]. However, since the energy is dissipated mostly in the approaching phase of normal motion, it may be more realistic to assume b_N dependent of the phase of normal motion, e.g., in the simple form:

$$\begin{aligned} b_N(\dot{a}) &= b_{N+} \quad \text{if } \dot{a} \geq 0 \\ b_N(\dot{a}) &= b_{N-} \quad \text{if } \dot{a} < 0. \end{aligned} \quad (3.32)$$

At this point it is important to emphasize that so far there exist no methods of evaluating, even approximately, the value of the coefficient b_N (l_N is usually assumed to be equal to m_N).

An attempt to estimate a reasonable range for the coefficient b_N was presented by Tworzydło and Becker [1988]. In this work some of these results will be used.

The next constitutive equation of the interface is the friction law, which in the applied model is of the form:

$$\begin{aligned} \text{if } a < 0 \text{ then } \sigma_T &= 0 \\ \text{if } a \geq 0 \text{ then } |\sigma_T| &\leq c_T(a)^{m_T} \text{ and} \end{aligned} \quad (3.33)$$

$$|\sigma_T| \leq c_T(a)^{m_T} \Rightarrow \dot{a} = 0$$

$$|\sigma_T| = c_T(a)^{m_T} \Rightarrow \dot{a} = -\lambda \sigma_T \dot{t} \quad (\lambda > 0)$$

where index T refers to the direction tangential to the contact surface.

It is important to observe that the friction force is the function of the normal approach of the two surfaces which, in turn, depends on the normal force. The actual value of the static coefficient of friction can be expressed in the form:

$$\mu = \frac{c_T}{c_N} a^{(m_T - m_N)} \quad (3.34)$$

so that if $m_T = m_N$, this model represents the Coulomb friction law with $\mu = \frac{c_T}{c_N}$.

In terms of normal stress on the interface the equation (3.34) can be recast in the form

$$\mu = \frac{c_T}{c_N^{\frac{m_T}{m_N}}} (-\sigma_N)^{\frac{m_T}{m_N} - 1} \quad (3.35)$$

where it was assumed that $\dot{a} = 0$ (static friction). From equations (3.34) and (3.35) one can easily observe that there are at least two ways of modeling load (or penetration) dependence of the coefficient of friction

1. by considering $m_T \neq m_N$. However, in this case μ at loads decreasing to zero will either decrease to zero (for $m_T < m_N$) or grow to infinity (for $m_T > m_N$).
2. by assuming $m_T = m_N$ and c_T dependent on normal load (or penetration). This approach seems to be more general and will be used further in this work.

Another important observation is that within this model the value of friction force is *not* uniquely defined in terms of displacements (in particular tangential displacements). In order to avoid difficulties resulting from that fact, the friction law is usually regularized. One of the possible regularized forms of friction law is (for two-dimensional problems):

$$\sigma_T = c_N a_+^{m_T} \Phi_\epsilon(\dot{d})$$

where

$$\Phi_\epsilon(\xi) = \begin{cases} \left(2 - \left|\frac{\xi}{\epsilon}\right|\right) \frac{\xi}{\epsilon} & \text{if } |\xi| \leq \epsilon \\ \text{sgn } \xi & \text{if } |\xi| > \epsilon \end{cases} \quad (3.36)$$

is the regularization function.

The details of regularization of the friction law are discussed elsewhere in this report (Chapter 2).

4 Variational Formulation of a Boundary-Value Problem

It is useful, for the sake of numerical solution, to recast the boundary problem formulated in the previous section in a weak, variational form. Since, as it was mentioned, friction force in (3.33) is not uniquely defined in terms of displacements, the corresponding variational formulation takes the form of variational inequality (see, e.g., Oden and Martins [1985]).

Kikuchi, Oden [1988], Panogiatopoulos [1975], Böhm [1985], Duvaut and Lions [1976], and many others). If, however, the regularized version of friction law (3.36) is used, then the variational formulation becomes an equality of a form similar to the principle of virtual power. For our model the variational problem with regularized friction law is:

Find $\mathbf{u}(t) : [0, T] \rightarrow V$ such that:

$$\begin{aligned} & \langle \mathbf{i}(\mathbf{a}(t)), \mathbf{w} \rangle + e(\mathbf{u}(t), \mathbf{w}) + \langle \mathbf{k}(\mathbf{u}(t)), \mathbf{w} \rangle + \\ & + \langle \mathbf{c}(\mathbf{v}(t)), \mathbf{w} \rangle + \langle \mathbf{p}(\mathbf{u}(t), \mathbf{v}(t)), \mathbf{w} \rangle + \\ & + \langle \mathbf{j}(\mathbf{u}(t), \mathbf{v}(t)), \mathbf{w} \rangle = \langle \mathbf{f}(t), \mathbf{w} \rangle \quad \forall \mathbf{w} \in V \end{aligned} \quad (4.1)$$

with initial conditions:

$$\begin{aligned} \mathbf{u}(0) &= \mathbf{u}_0 \quad \text{in } \Omega \\ \mathbf{v}(0) &= \mathbf{v}_0 \quad \text{in } \Omega \end{aligned} \quad (4.2)$$

In the above formula V is the space of admissible displacements $V = \{\mathbf{w} \in \mathbf{H}^1(\Omega), \gamma_0(\mathbf{w}) = \mathbf{0} \text{ on } \Gamma_D, w_N \stackrel{\text{not}}{=} \gamma_N(\mathbf{w}) = 0 \text{ on } \Gamma_N^F\}$ where $\Omega = \Omega_A \cup \Omega_F$, γ_0 and γ_N are relevant trace maps and the symbol $\langle \cdot, \cdot \rangle$ represents duality pairing on $V' \times V$.

The exact form and meaning of consecutive terms in (4.1) will be explained below. For the sake of clarity we will not distinguish bodies by extra superscripts — this assignment is defined by the domain of integration. The consecutive terms denote:

$$\langle \mathbf{i}(\mathbf{a}), \mathbf{w} \rangle = \int_{\Omega_A} \rho \mathbf{a} \cdot \mathbf{w} dx + \int_{\Omega_F} \rho \mathbf{a}^B \cdot \mathbf{w} dx$$

– virtual power of inertia forces, where \mathbf{a}^B is expressed by (3.15),

$$e(\mathbf{u}, \mathbf{w}) = \int_{\Omega} \mathbf{T} \cdot (\nabla \mathbf{w}) dx$$

– virtual power of internal stresses,

$$\langle \mathbf{k}(\mathbf{u}), \mathbf{w} \rangle = \int_{\Gamma_E^A} \mathbf{K}_E \mathbf{u} \cdot \mathbf{w} ds + \int_{\Gamma_E^F} k u_N w_N ds$$

– virtual power of supporting springs,

$$\langle \mathbf{c}(\mathbf{v}), \mathbf{w} \rangle = \int_{\Gamma_D^F} c v_N w_N ds$$

– virtual power of supporting dashpots,

$$\begin{aligned} \langle \mathbf{p}(\mathbf{u}, \mathbf{v}), \mathbf{w} \rangle &= \int_{\Gamma_c} \sigma_N(\mathbf{u}, \mathbf{v}) w_N^R ds \\ &= \int_{\Gamma_c^F} \sigma_N(\mathbf{u}, \mathbf{v}) w_N^F dA - \int_{\Gamma_c^A} \sigma_N(\mathbf{u}, \mathbf{v}) w_N^A ds \end{aligned}$$

– virtual power of contact normal forces on the interface.

$$\langle j(\mathbf{u}, \mathbf{v}), \mathbf{w} \rangle = - \int_{\Gamma_c^F} \partial_T(\mathbf{u}, \mathbf{v}) w_T^F ds - \int_{\Gamma_c^A} \partial_T(\mathbf{u}, \mathbf{v}) w_T^A ds$$

– virtual power of friction forces on the interface

$$\langle \mathbf{p}(t), \mathbf{w} \rangle = \int_{\Omega} \mathbf{b} \cdot \mathbf{w} dx + \int_{\Gamma_F} \mathbf{p} \cdot \mathbf{w} ds$$

– virtual power of external loads.

The symbols used in the above formulas are defined in the preceding sections.

It can be shown by standard procedure that the above problem is — for sufficiently regular fields — equivalent to the original boundary-value problem with regularized friction.

The question of existence and uniqueness of a solution to the problem (4.1) will not be addressed in this work, and we will focus on a numerical solution of its described form.

5 Discretization of the Boundary Value Problem

5.1 Finite Element Discretization

For the sake of solving the variational problem (4.1) using a finite dimensional model, we approximate displacements, velocities and accelerations by the functions from space $V^h \subset V$, expressed in the form:

$$\begin{aligned} \mathbf{u}^h(\mathbf{x}, t) &= \sum_{I=1}^N \mathbf{u}^I(t) \Psi^I(\mathbf{x}) \\ \mathbf{v}^h(\mathbf{x}, t) &= \sum_{J=1}^N \mathbf{v}^J(t) \Psi^J(\mathbf{x}) \\ \mathbf{a}^h(\mathbf{x}, t) &= \sum_{J=1}^N \mathbf{a}^J(t) \Psi^J(\mathbf{x}) \end{aligned} \quad (5.1)$$

where N is the number of discretization nodes in the domain, $\mathbf{u}^h(\mathbf{x}, t) \dots \mathbf{a}^h(\mathbf{x}, t)$ are displacements, velocities and accelerations of nodes, and $\Psi^I(\mathbf{x})$ is the shape function associated with node I .

For such defined approximation of displacements the discretized version of (4.1) is of the form:

$$\begin{aligned} &\text{Find } \mathbf{u}^h(t) : [0, T] \rightarrow V^h \text{ such that} \\ &< \mathbf{i}(\mathbf{a}^h(t)), \mathbf{w}^h > + e(\mathbf{u}^h(t), \mathbf{w}^h) + < \mathbf{k}(\mathbf{u}^h(t)), \mathbf{w}^h > + \\ &+ < \mathbf{c}(\mathbf{v}^h(t)), \mathbf{w}^h > + < \mathbf{p}(\mathbf{u}^h(t), \mathbf{v}^h(t)), \mathbf{w}^h > + \\ &+ < \mathbf{j}(\mathbf{u}^h(t), \mathbf{v}^h(t)), \mathbf{w}^h > = < \mathbf{f}(t), \mathbf{w}^h > \quad \forall \mathbf{w}^h \in V^h \end{aligned} \quad (5.2)$$

with initial conditions:

$$\begin{aligned} \mathbf{u}^h(0) &= \mathbf{u}_0^h \quad \text{in } \Omega^h \\ \mathbf{v}^h(0) &= \mathbf{v}_0^h \quad \text{in } \Omega^h \end{aligned} \quad (5.3)$$

By the standard procedure, namely using divergence theorem in $a(\mathbf{u}^h(t), \mathbf{w}^h)$ and requiring that (5.2) be satisfied for every \mathbf{w}^h , we obtain equivalent form of (5.2):

$$M\ddot{\mathbf{R}}(t) + (C + \widehat{C})\dot{\mathbf{R}}(t) + (K + \widehat{K})\mathbf{R}(t) + \quad (5.4)$$

$$\mathbf{P}(\dot{\mathbf{R}}(t), \mathbf{R}(t)) + \mathbf{J}(\dot{\mathbf{R}}(t), \dot{\mathbf{R}}(t)) = \mathbf{F}(t) \quad (5.5)$$

with initial condition:

$$\begin{aligned} R(0) &= R_0 \\ \dot{R}(0) &= \dot{R}_0 \end{aligned} \quad (5.6)$$

In the above formula:

- R, \dot{R}, \ddot{R} the column vector of nodal displacements, velocities and accelerations,
 M, D, K standard mass, damping and stiffness matrices.
 \widehat{C}, \widehat{K} damping and stiffness matrices resulting from additional inertia terms for frame F ,
 F consistent nodal load vector,
 P, J vector of consistent nodal forces due to normal and frictional response of the interface, respectively.

The components of vectors P and J for each element can be calculated from:

$$\begin{aligned} {}^{(e)}P_{Ij}^\Delta &= - \int_{(\epsilon)\Gamma_c} \sigma_N n_j^A \Psi_I^\Delta i(\Delta) ds \quad \Delta = A, F \\ {}^{(e)}J_{Ij}^\Delta &= - \int_{(\epsilon)\Gamma_c} \sigma_T t_j \Psi_J^\Delta i(\Delta) ds \quad \Delta = A, F \end{aligned} \quad (5.7)$$

where

$$i(\Delta) = \begin{cases} 1 & \text{if } \Delta = A \\ -1 & \text{if } \Delta = F \end{cases} \quad (5.8)$$

while components of additional element matrices ${}^{(e)}\widehat{C}$, ${}^{(e)}\widehat{K}$ for elements covering frame F are given by

$$\begin{aligned} {}^{(e)}\widehat{C}_{IJKl} &= 2 \int_{(\epsilon)\Omega_F} \rho \Psi_I \Psi_{J,m} \dot{U}_m \delta_{kl} d\Omega_e \\ {}^{(e)}\widehat{K}_{IJKl} &= \int_{(\epsilon)\Omega_F} \rho \Psi_I \Psi_{J,mn} \dot{U}_m \dot{U}_n \delta_{kl} d\Omega_e \\ &= - \int_{(\epsilon)\Omega_F} \rho \Psi_{I,n} \Psi_{J,m} \dot{U}_m \dot{U}_n \delta_{kl} d\Omega_e \end{aligned} \quad (5.9)$$

It is of interest to note that the matrix \widehat{C} is *nonsymmetric*.

5.2 Incremental Equations of Motion

The semidiscrete equations of motion, presented in the previous section, are—due to non-linearity of response of the interface—generally nonlinear.

However, for the sake of analysis of infinitesimally small vibrations around some prescribed configuration of the system, defined by $\mathbf{R}_R = \mathbf{0}$, $\ddot{\mathbf{R}}_R = \mathbf{0}$, one can use the linearized version of 5.5:

$$\mathbf{M}\ddot{\mathbf{R}} + (\mathbf{C} + \widehat{\mathbf{C}} + \mathbf{C}^N + \mathbf{C}^T)\dot{\mathbf{R}} + (\mathbf{K} + \widehat{\mathbf{K}} + \mathbf{K}^N + \mathbf{K}^T)(\mathbf{R} - \mathbf{R}_R) = \Delta \mathbf{F} \quad (5.10)$$

where $\mathbf{R} - \mathbf{R}_R$ is the infinitesimal increment of displacements, \mathbf{M} , \mathbf{C} , $\widehat{\mathbf{C}}$, $\widehat{\mathbf{K}}$ have the same meaning as in 5.5 and matrices \mathbf{K}^N , \mathbf{K}^T , \mathbf{C}^N , \mathbf{C}^T can be calculated as derivatives at \mathbf{R}_R :

$$\begin{aligned} \mathbf{K}^N &= \frac{\partial \mathbf{P}}{\partial \mathbf{R}} \\ \mathbf{K}^T &= \frac{\partial \mathbf{J}}{\partial \mathbf{R}} \\ \mathbf{C}^N &= \frac{\partial \mathbf{P}}{\partial \dot{\mathbf{R}}} \\ \mathbf{C}^T &= \frac{\partial \mathbf{J}}{\partial \dot{\mathbf{R}}} \end{aligned} \quad (5.11)$$

with the use of formulas presented in Section 4.

5.3 Rigid Body Model

In our numerical analysis of an experimental setup, which will be presented in Part II of this paper, we will use a relatively simple model consisting of three rigid bodies connected with elastic springs and linear dashpots. The use of such a model will be justified in Part II, where we will briefly present forms of equations of motion (5.5) and (5.10).

The model, presented in Fig. 7 and representing experimental setup shown in Fig. 8 consists of two rigid bodies, A and C, and a rigid frame F. Body A corresponds to the pin, body C represents the heavy block on which the pin was mounted, and rigid frame F represents the moving disk (body B). The dimensions of each body are length L and height H.

Each of the bodies has three degrees of freedom: displacements of center of mass U_x, U_y and counterclockwise rotation θ .

Frame F and body C are supported by elastic springs of stiffnesses $\overset{\delta}{K}_x, \overset{\delta}{K}_y, \overset{\delta}{K}_\theta$, $\Delta = F, C$, and body A is connected with body C by means of three springs $K_x^{A-C}, K_y^{A-C}, K_\theta^{A-C}$. For each of these springs there may exist a corresponding dashpot.

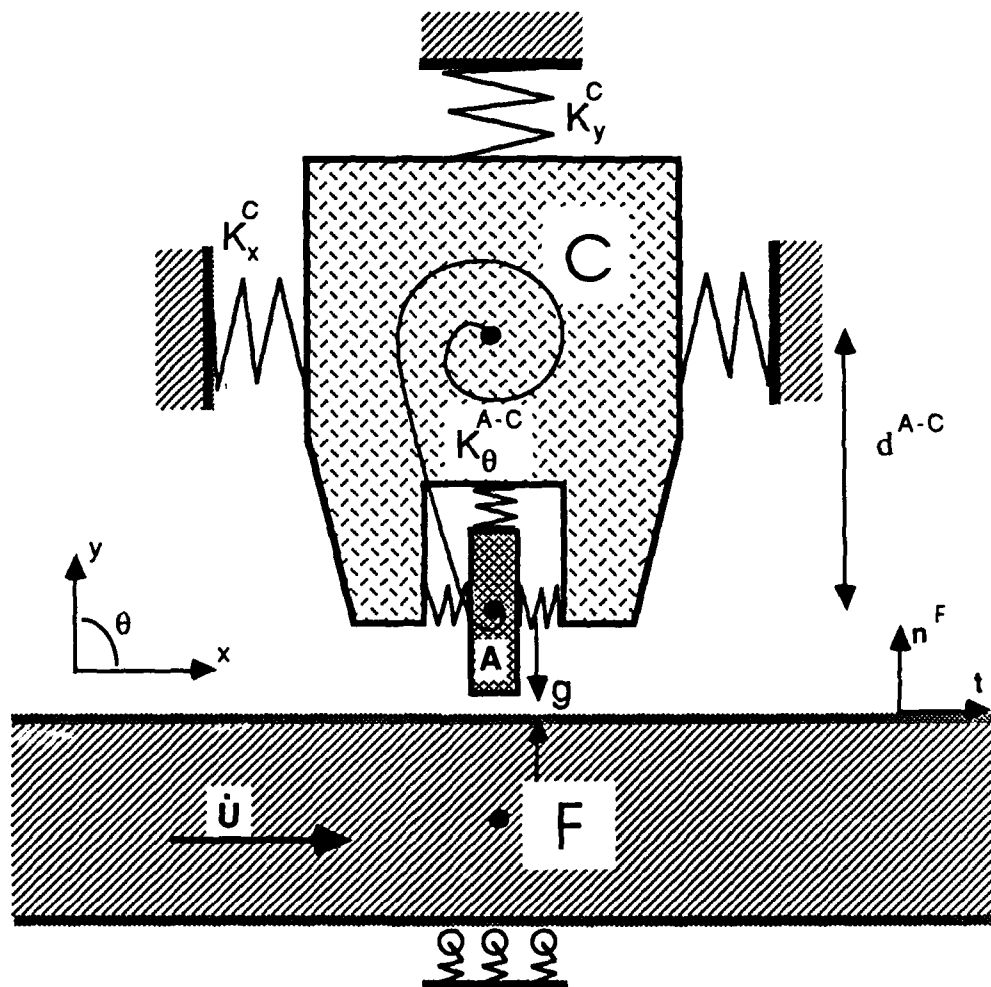


Figure 7: Rigid body model of an experimental pin-on-disk apparatus.

Regarding kinematics of the model, we will assume, in agreement with the general model presented in Section 3 that displacements are basically infinitesimal, with a possible large translation of assembly $A \cup C$ along the disk F .

The initial configuration of our model is presented in Fig. 7. However, in order to model different angles of attack, we will permit some initial rotation $\bar{\theta}$ of frame F . This rotation is assumed to be moderate in the sense that $\cos \bar{\theta} \approx 1$ and $\sin \bar{\theta} \approx \bar{\theta}$.

Since the original pin used in the experiment was of the circular cross-section, we allow variable distribution of thickness of the pin $b^A(s)$, $s \in [-L^A/2, L^A/2]$ (this is important in calculations of response of the interface).

The approximation of displacements for each rigid body can formally be written in the form similar to (5.1).

$$\begin{aligned} u_x(X, Y, t) &= U_x(t) + (\cos \theta(t) - 1)(X - \dot{X}) + \sin \theta(t)(Y - \dot{Y}) \\ u_y(X, Y, t) &= U_y(t) - \sin \theta(t)(X - \dot{X}) + (\cos \theta(t) - 1)(Y - \dot{Y}) \end{aligned} \quad (5.12)$$

where $u_x(X, Y, t)$ is the displacement of any point in the body, $U_x(t), U_y(t), \theta(t)$ are discrete displacements defined in the center of mass, and \dot{X}, \dot{Y} is the reference position of the center of mass (of the given rigid body). Approximation of velocities and accelerations can be obtained by appropriate time differentiation of the above formula.

For such defined discretizations the equations of motion are of exactly the same form (5.5) as for the finite element model, with the displacement vector

$$\mathbf{R} = \{U_x^A, U_y^A, \theta^A, \dots, \theta^C\}^T. \quad (5.13)$$

and loads:

$$\mathbf{F} = \{F^A, F_y^A, F_\theta^A, \dots, F_\theta^C\}^T. \quad (5.14)$$

The mass matrix for this model is of the diagonal form:

$$\mathbf{M} = [M^A, M^A, I^A, \dots, I^C] \quad (5.15)$$

and the stiffness matrix \mathbf{K} is:

$$\begin{bmatrix}
K_x^{A-C} & 0 & 0 & 0 & 0 & 0 & -K_x^{A-C} & 0 & -DK_x^{A-C} \\
0 & K_y^{A-C} & 0 & 0 & 0 & 0 & 0 & -K_y^{A-C} & 0 \\
0 & 0 & K_\theta^{A-C} & 0 & 0 & 0 & 0 & 0 & -K_\theta^{A-C} \\
0 & 0 & 0 & K_x^F & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & K_y^F & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & K_\theta^F & 0 & 0 & 0 \\
-K_x^{A-C} & 0 & 0 & 0 & 0 & 0 & K_{77} & 0 & DK_x^{A-C} \\
0 & -K_y^{A-C} & 0 & 0 & 0 & 0 & 0 & K_{88} & 0 \\
-DK_x^{A-C} & 0 & -K_\theta^{A-C} & 0 & 0 & 0 & DK_x^{A-C} & 0 & K_{99}
\end{bmatrix} \quad (5.16)$$

where $D = d^{A-C}$ and $K_{77} = K_x^C + K_x^{A-C}$, $K_{88} = K_y^C + K_y^{A-C}$, $K_{99} = K_\theta^C + K_\theta^{A-C} + K_x^{A-C} D^2$.

Since the body B is rigid, the additional stiffness matrix \hat{K} is zero and the additional damping matrix has only two non-zero elements $C_{56} = 2\dot{U}M^F$, $C_{46} = 2\dot{U}M^F\theta^F$. The two vectors P and I , representing normal and frictional response of the interface, are of the form:

$$\begin{aligned}
P &= \{P_x^A, P_y^A, P_\theta^A, P_x^F, P_y^F, P_\theta^F, 0, 0, 0\} \\
I &= \{I_x^A, I_y^A, I_\theta^A, I_x^F, I_y^F, I_\theta^F, 0, 0, 0\}
\end{aligned} \quad (5.17)$$

where the consecutive components can be expressed in the form

$$\begin{aligned}
P_y^A &= \int_{\Gamma_c} b^A(s) [c_N a(s)^{m_N} + b_N(\dot{a}) a(s)^{l_N} \dot{a}(s)] ds \\
P_x^A &= -P_y^A \cdot \theta^B \\
P_\theta^A &= \int b^A(s) [c_N a(s)^{m_N} s + b_N(\dot{a}) a(s)^{l_N} \dot{a}(s) s] ds \\
P_i^F &= -P_i^A \quad i = x, y, \theta \\
J^A &= -\Phi_c(\dot{d}) \int_{\Gamma_c} b^A(s) c_T(a) a(s)^{m_T} ds \\
J_y^A &= J_x^A \cdot \theta^B \\
J_0^A &= J_x^A \cdot \frac{H^A}{2} \\
J_i^F &= -J_i^A \quad i = x, y, \theta
\end{aligned} \quad (5.18)$$

In the above formulas s spans the "bottom" of the pin and Γ_C^A is the line of actual contact between pin and disk, determined by $a(s) > 0$. Approach a can be calculated from a formula consistent with (3.24):

$$a(s) = U_y^F - U_y^A - s\theta^A + s^F\theta^F \quad (5.19)$$

where

$$s^F(s) = U_x^A - U_x^F + \theta^A \frac{H^A}{2} + \theta^F \frac{H^F}{2} + s \quad (5.20)$$

and the sliding velocity on the interface is given by the formula consistent with (3.29):

$$\dot{d} = \dot{U}_x^B + \dot{U}_x^A + \frac{H^F}{2}\dot{\theta}^F - \frac{H^A}{2}\dot{\theta}^A + \dot{U} \quad (5.21)$$

The linearized equations of motion for the rigid body model are of the same form (5.10) as for the Finite Element model. The matrices K^N, K^T, C^N, C^T occurring in this equation can be obtained by simple differentiation of the above formulas according to (5.11).

5.4 Remarks on Methods of Solving The Equations of Motion

In our approach the frictional behavior of the sliding system is modeled via transient solution of the system of equations of motion (5.5) or, in some cases, via analysis of natural frequencies of linearized equations (5.10).

The primary method of transient analysis was the Newmark Method, with the adaptive time increment applied previously by Oden, Martins [1985]. In some cases, primarily for verification purposes, the standard Gears or Adams integrators were applied (general purpose library routines were used). The natural frequencies of the linearized system (5.10) were calculated via the solution of the eigenproblem:

$$M\lambda^2 + (C + \widehat{C} + C^N + C^T)\lambda + (K + \widehat{K} + K^N + K^T) = 0 \quad (5.22)$$

which was obtained by the Householder method. The eigenvalues and eigenvectors of this system are, due to nonsymmetry of matrices, complex variables.

Part II

Analysis of Frictional Behavior of Pin-on-Disk Apparatus

6 General Presentation of Numerical Examples

In the following few sections we will apply the model developed in the first part of the report to the analysis of frictional behavior—in particular, stability—of a typical pin-on-disk setup.

As the particular object of our analysis we have chosen the apparatus applied by Aronov, D'Souza, Kalpakjian and Shareef [1984] in their experiment on frictional vibrations. The motivations for this choice were the documentation of both their experimental setup and results, and the relative simplicity of the applied apparatus as well as interesting results regarding self-excited oscillations and evolution of the kinetic coefficient of friction at increasing loads.

The motivation of our analysis is to further develop understanding of the nature of dynamic friction and to verify the applicability of the model presented in Part I to modeling of behavior of frictional systems.

In particular we would like to:

- identify the most important parameters affecting the occurrence of self-excited oscillations
- study the nature of motion of the system in the self-excited zone
- model the difference between static and kinetic coefficients of friction
- identify further directions of development of models of friction and relevant computational methods

It is important to point out that, although our basic reference experiment performed by Aronov, et. al [1984] was very well documented, still some data necessary in numerical modeling are missing and were not measured at the time of the experiment. That is why in some cases we used information presented in other works (e.g., constitutive constants for the interface). Moreover, we often “perturbed” the data from the basic data set, usually in order to better illustrate specific features of frictional behavior of the system under consideration.

As it was mentioned, the understanding of mechanisms of frictional stability and difference between static and kinetic friction is still far from being complete. In fact, one of the primary goals of this paper is to further develop our understanding of these phenomena and the most important parameters affecting them.

That is why we have decided to use as simple as possible—but still representative—a mechanical model of the system. This model, consisting of several rigid elements, was presented in the previous section. However, since this model can be considered as a certain variant of discretization of the problem, we can cautiously generalize our observations to more complex mechanical systems and other discretizations, in particular Finite Element models.

Our analysis consists of several stages. First we analyze the stability of frictional sliding at a wide range of loads. The stability and occurrence of self-excited oscillations is estimated by means of eigenvalue analysis of linearized equations of motion (5.10) and the results are confirmed by a transient analysis of full nonlinear equations of motion (5.5).

This transient analysis was also the basis for the modeling of the kinetic coefficient of friction, in particular drop of friction force in the zone of self-excited oscillations.

In the last sections we have extended our analysis to include angles of attack different from zero angles of attack. This analysis—which is beyond the actual program of the original experiment—gives a very interesting conclusion regarding sensitivity of the nature of frictional sliding to the angle of attack.

7 Summary of Experimental Results

The main goal of our numerical analysis was to model frictional behavior of an experimental setup, described by Aronov, D'Souza, Kalpakjian and Shareef [1984]. The detailed description of the apparatus is presented by these authors in the above referenced paper. Here we will briefly present the basic components and simplified sketch of the experimental setup (Fig. 8).

The apparatus consists of a rotating iron disk (1), 200mm in diameter, and a steel pin (2), 5mm in diameter. The pin is mounted on an elastic arm (3), fixed at one end. A heavy, rigid block (4) was fixed at the same end as the pin. It is of importance that the pin was not really fixed to the supporting arm. A force transducer was sandwiched between the pin and the supporting arm, so this connection did actually have some compliance.

During the experiment the normal load was exerted on the pin by applying a moment at the fixed end of the supporting arm. The iron disk was rotating at constant angular velocity, such that the average sliding speed of the pin was equal $\dot{U} = 46\text{cm/s}$. The part of the experiment which was of greatest interest to us consisted of analysis of evolution of the kinetic coefficient of friction and occurrence of self-excited oscillations at normal loads increasing from 0 to 200 N.

The experimental results are shown in 9, which is a simplified version of Fig.9 by Aronov, et al. [1984]. The figure presents a plot of the kinetic coefficient of friction versus normal load on the pin. There are four general regions on this plot: (for detailed description refer

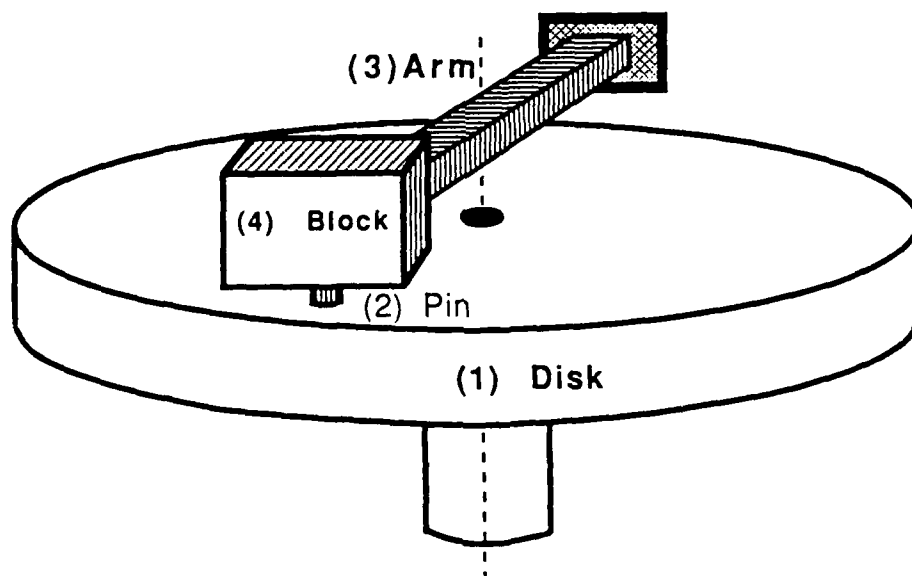


Figure 8: Experimental apparatus used by Aronov et al. [1985]

to Aronov, et al. [1984].

- I. Region of steady-state friction with very small vibrations caused by interlocking of imperfections on the interface
- II. Region of nonlinear friction with small vibrations
- III. Transient region with periodic bursts of self-excited oscillations accompanied by drop of the apparent kinetic coefficient of friction
- IV. Region of self-excited oscillations, i.e., high-amplitude oscillations perceived as a frictional noise. In this region the kinetic coefficient of friction was lower than in other zones.

It is of importance that the coefficient of friction plotted in Fig. 9 was not measured on the interface, but was estimated by measurements of deformation of the end of the supporting arm. Therefore it can be considered to be an apparent or macroscopic coefficient of friction.

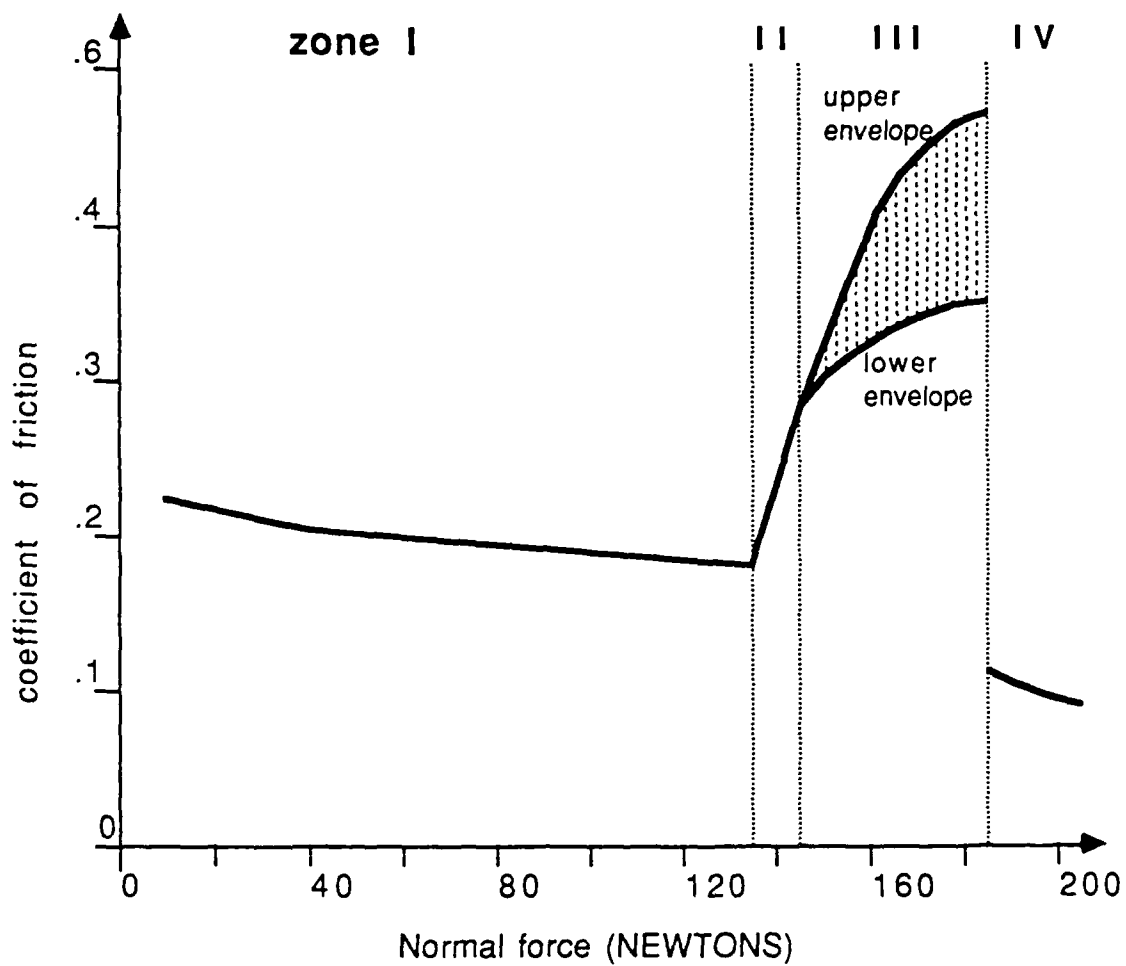


Figure 9: Schematic plot of experimental results—apparent coefficient of friction vs. normal load.

8 Analysis of Stability of Frictional Sliding

The occurrence of self-excited oscillations is the consequence of dynamic instability of the system, that is, the situation in which any perturbation of steady-sliding causes growth of oscillations and eventually occurrence of high-amplitude vibrations. This observation was made by Aronov, et al. [1984] and modeled numerically for a simple test model by Oden and Martins [1985]. However, these researchers did not specify the reasons or circumstances that can create such situations. This problem was further studied by Becker and Tworzydło [1988] and their general observation was that:

The self-excited oscillations occur when the natural frequencies of normal and rotational vibrations of the slider (in contact with a slide line) are relatively close to one another. Then the existence of friction causes coupling between rotational and normal modes and, as a consequence, frequency coalescence and propagation of self-excited oscillations.

The above observation will be discussed in more detail in the following sections. At this point it is important to observe that the crucial parameters in the analysis of stability are all parameters affecting normal and rotational vibrations of the system.

If one looks at the model of experimental apparatus presented in Fig. 8 one can observe that the frequency of rotational oscillations of the slider depends primarily on the torsional stiffness of the supporting arm. On the other hand, frequencies of normal vibrations depend strongly on the stiffness of all elements between the slider (rigid block in Fig. 8) and the disk. Since the frictional interface is generally quite stiff, the compliance of connection between pin and block becomes extremely important. That is why in the computational model presented in Fig. 7 linear springs were introduced between body A (pin) and body C (block). These springs represent the compliance of force transducers and other elements between the pin and the block. It should be pointed out that this is a simplified approximation, since the characteristics of the real connection was probably strongly nonlinear. This characteristic was, of course, not known in the experiment (it was not considered important and, moreover, it is very difficult to measure) and it will be one of primary parameters estimated and studied in this work.

8.1 Basic Computational Model

The model of experimental setup used in our computations was already presented in Section 3, Fig. 7. The basic data corresponding to the parameters of the experimental apparatus

are (notation corresponds to that of Section 3):

1. Body A (pin):

$$\begin{aligned} L &= 0.5 \text{ cm} && (\text{circular cross section}) \\ H &= 1.5 \text{ cm} \\ M &= 2.2972 \cdot 10^{-3} \text{ kg} \\ J_{\theta} &= 4.6663 \cdot 10^{-4} \text{ kg} \end{aligned}$$

2. Body C (block):

$$\begin{aligned} M^C &= 4.5 \text{ kg} \\ J_{\theta}^C &= 96.0 \text{ kg cm}^2 \\ d^{A-C} &= 3.085 \text{ cm} \\ K_x^C &= 845.6 \cdot 10^2 \text{ kg s}^{-2} \\ K_y^C &= 773.9 \cdot 10^3 \text{ kg s}^{-2} \\ K_{\theta}^C &= 432.5 \cdot 10^5 \text{ kg s}^{-2} \text{ cm}^{-2} \end{aligned} \tag{8.1}$$

3. Body F (disk):

$$\begin{aligned} L^F &= 20 \text{ cm} \\ H^F &= 8 \text{ cm} \\ M^F &= 10.0 \text{ kg} \\ M^F &= 24.0 \text{ kg} \\ J_{\theta}^F &= \text{BIG} \\ K_x^F &= K_y^F = K_{\theta}^F = \text{BIG} \end{aligned}$$

In the above data BIG represents a big number so that the support of the disk (shaft, bearings) is considered—in the basic data set—to be rigid. This was assumed at this stage due to the lack of experimental data, but in further analysis this assumption is relaxed.

The material constants for the interface were calculated from the table presented by Back, Burdekin and Cowley [1973] for the surface finish corresponding to that of disk and pin. These data are:

$$\begin{aligned} c_N &= 1.25 \text{ kg cm}^{3.5} \text{ s}^{-2} \\ c_T &= 0.3125 \cdot 10^3 \text{ kg cm}^{3.5} \text{ s}^{-2} \\ m_T &= 2.5 \\ m_N &= 2.5 \\ b_N &= l_N = 0 \end{aligned} \tag{8.2}$$

and they correspond to the coefficient of friction $\mu = 0.25$ (see equation (3.34)). The stiffness of connection between pin and block is defined in the forthcoming sections. It can be noted that, in this basic set of data, there is no damping in the system whatsoever.

8.2 Estimation of Stability via Eigenvalue Analysis

8.2.1 The Basic Algorithm

In the analysis of stability of frictional sliding we follow the general procedure used by Oden and Martins [1985]. This procedure consists of two stages:

1. Assume that there are no vibrations of the slider (steady-sliding) and find the equilibrium position in this state.
2. Perturb this equilibrium position and analyze dynamic stability of motion of the system.

Solution of the first stage is equivalent to solving the equation of motion (5.5) under the constraint, that velocities and accelerations are zero (except for prescribed velocity of the disk \dot{U}). Then (5.5) reduces to the quasistatic nonlinear equation:

$$(K + \hat{K}) + P(0, R) + J(0, R) = F \quad (8.3)$$

which can be solved by the Newton method. The solution of this equation corresponds to the steady-sliding equilibrium position.

This situation is, of course, idealized, since interlocking of the surface asperities provides permanent small perturbations of this equilibrium position. If the system is dynamically stable, vibrations caused by this interlocking do not propagate but stay at a low level, perceived as a steady sliding (zone I in Fig.9). If the system is, however, dynamically unstable, after perturbation of the steady-sliding equilibrium position the oscillations grow and reach high values corresponding to zone IV in Fig. 9.

The stability of motion of the slider after perturbation of the steady-sliding equilibrium position can be estimated by transient analysis of (5.5) or, in a simplified version, by analysis of the stability of linearized equations of motion (5.10).

The latter can be reduced to the solution of the eigenvalue problem:

$$M\lambda^2 + (C + \hat{C} + C^N + C^t)\lambda + (K + \hat{K} + K^N + K^T) = 0 \quad (8.4)$$

Due to existence of damping and asymmetry of the matrix $[K + K^N + K^T]$ the eigenvalues of this problem are, in general, complex. The dynamic stability of the system is determined by the values of real parts of these eigenvalues. In general there are three situations possible:

1. $Re(\lambda_i) = 0$ for $i = 1 \dots n$ - we can expect stable oscillations about the static equilibrium position,
2. $Re(\lambda_i) < 0$ for $i = 1 \dots n$ - we can expect stable oscillations which are gradually damped out,

3. $\exists i$ s.t. $\text{Re}(\lambda_i) > 0$ - we can expect unstable, growing oscillations.

It is important that this analysis of eigenvalues establishes only the stability of the linearized small oscillations about the equilibrium position. This gives only the initial estimate of the behavior of the model, which is essentially nonlinear.

8.2.2 Influence of Normal Stiffness on the Stability of Frictional Sliding

As stated before, the key factor in the occurrence of self-excited oscillations of the slider (understood here as a pin-block assembly) is the frequency of its normal vibrations when the pin is in contact with the disk. The self-excited oscillations occur when this frequency is close to the frequency of rotational oscillations which, for the model under consideration, is determined by the stiffness of supporting arm K_θ^C and is almost constant under all loads (about 126 Hz). The normal stiffness of the system depends on the stiffness of the supporting arm K_y^C , stiffness of the interface, and of the spring K_y^{A-C} . The last value was not determined in the experiment and, as a matter of fact, it was not considered important. Therefore one of the goals of this work is to point out the importance of this factor and study its influence on the occurrence of self-excited oscillations.

Before performing this study it is necessary to show how, for given parameters of the system, the range of instability zone is established.

As a representative example we select the basic data set presented in Section 8.1 with the normal stiffness $K_y^{A-C} = 2.9 \cdot 10^6 \text{ kg s}^{-2}$ and other springs assumed temporarily to be rigid: $K_x^{A-C} = K_\theta^{A-C} = \text{BIG}$. If we consider this system *without* friction, then the dependence of normal and rotational frequencies on the normal load is represented by dashed lines on Fig. 10. These two curves intersect at the value of normal load $F_N = 24 \text{ N}$. Introduction of friction ($\mu = 0.25$) into the system provides coupling between normal and rotational modes and coalescence of corresponding frequencies in the vicinity of intersection points (solid lines on Fig. 10). The corresponding eigenvalues have positive real parts, so this zone is the region of dynamic instability and, as a consequence, self-excited oscillations. The width of this instability zone increases with increasing value of the coefficient of friction.

The above analysis was performed for various stiffnesses K_y^{A-C} and thus the general range of unstable zone was established. This zone is presented in a logarithmic scale in Fig. 11. It can be observed that if the compliance of all elements between pin and rigid block is disregarded ($K_y^{A-C} \rightarrow \infty$), then instabilities occur at very small values of normal forces, quite different than in the experiment.

On the other hand, at $K_y^{A-C} = 2.057 \cdot 10^6 \text{ kg s}^{-2}$ the self-excited oscillations occur at loads greater than 185 N, which corresponds well to experimental observations. This value of K_y^{A-C} was therefore included in the basic data set for the model of experimental setup.

The detailed plot of frequencies of normal and rotational vibrations in this case is presented in Fig. 12. It can be observed that the width of instability zone depends strongly

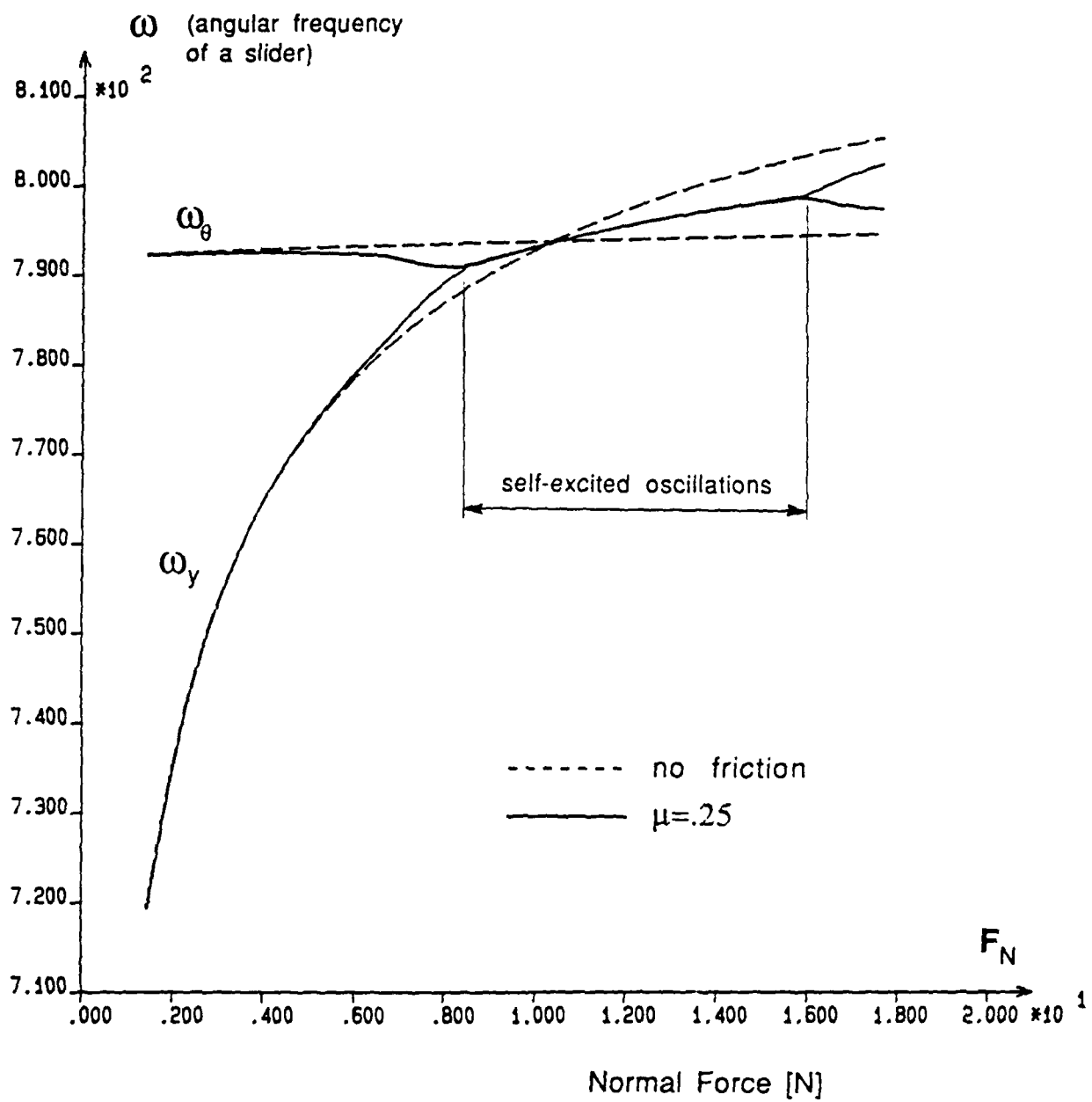


Figure 10: Estimation of instability zone via analysis of natural frequencies.
 $K_y^{A-C} = 2.4 \cdot 10^6 \text{ kg s}^{-2}$

on the coefficient of friction, and perturbation of μ from 0.25 to 0.30 reduces normal force corresponding to onset of self-excited oscillations to $120N$.

This sensitivity of the range of self-excited vibrations to the frictional properties of the interface can explain existence—in the experimental observations—of the region of transient friction with periodic bursts of self-excited oscillations (zone III in Fig. 11). According to experimental researchers in this zone “the mean value of friction force increases with time at a constant normal load” and “when the mean friction force reached a sufficiently high value, a temporary burst of self-excited vibrations would occur and the friction force would instantly fall to a lower value” (Aronov, et al. [1984]).

This observation is in perfect accord with sensitivity of the range of instability zone to the coefficient of friction, presented in Fig. 12.

8.2.3 Influence of Other Parameters on the Frictional Stability

The normal compliance of a pin-to-block connection is, as it was shown, a crucial factor in the occurrence of self-excited oscillations. In this paragraph we will present analysis of influence of some other parameters of the system on the stability of frictional sliding. However, we did not analyze parameters, which role is relatively easy to anticipate (like stiffness of the supporting arm), but we focused our attention on selected parameters, which give not so easily predictable results. In particular, analyzed were:

- stiffnesses K_x^{A-C} and K_θ^{A-C} of a pin-to-block connection,
- stiffness of the disk support.

From the numerical study which will not be presented here, the conclusion can be drawn, that horizontal and rotational stiffnesses of a pin-to-block connection have virtually no effect on natural frequencies of vibrations of the slider (pin-block assembly) and, therefore, on its frictional stability. However, at certain values of stiffness K_θ^{A-C} the natural normal and rotational frequencies of the pin itself can be close enough to permit self-excited, high frequency oscillations of the pin. For example, for general data given in section 8.1 and:

$$\begin{aligned} K_y^{A-C} &= 2.4 \times 10^6 \text{ kg s}^{-2} \\ K_\theta^{A-C} &= 2.4 \times 10^6 \text{ kg cm}^2 \text{ s}^{-2} \\ \mu &= 0.25 \end{aligned} \tag{8.5}$$

the plot of rotational and normal frequencies of both the pin and the block are presented in Fig. 13.

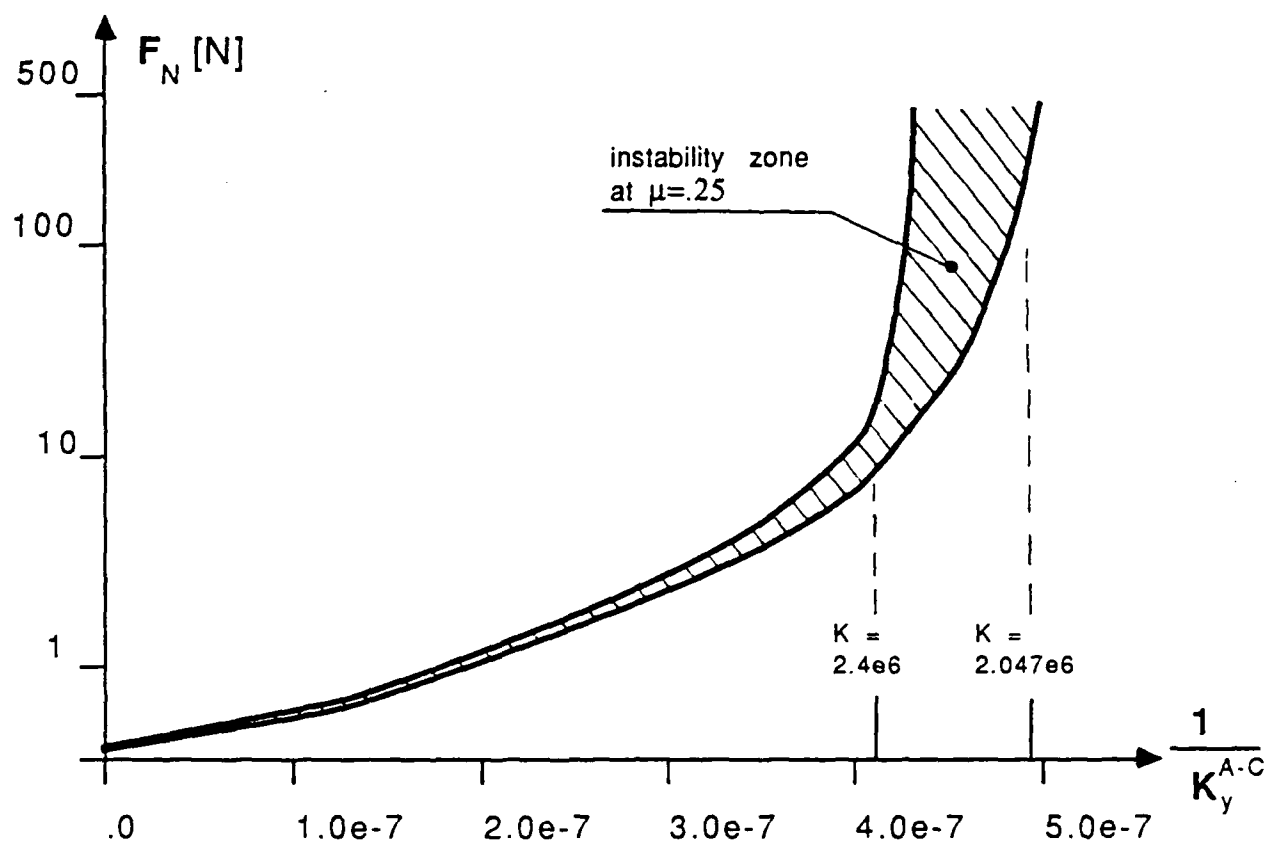


Figure 11: Range of frictional instability at various values of K_y^{A-C}

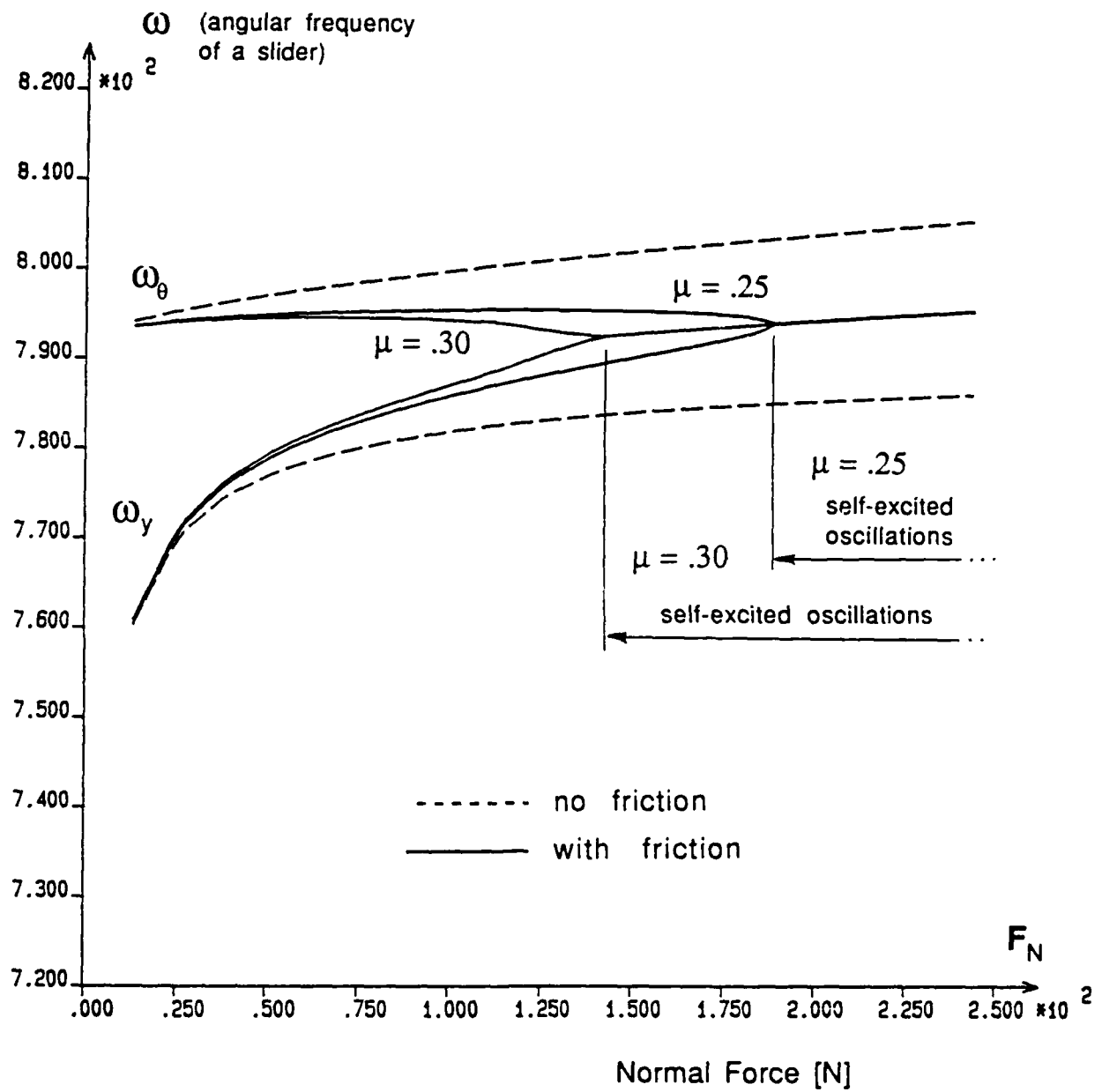


Figure 12: Estimation of instability via analysis of natural frequencies,
 $K_v^{A-C} = 2.057 \text{ kgs}^{-2}$

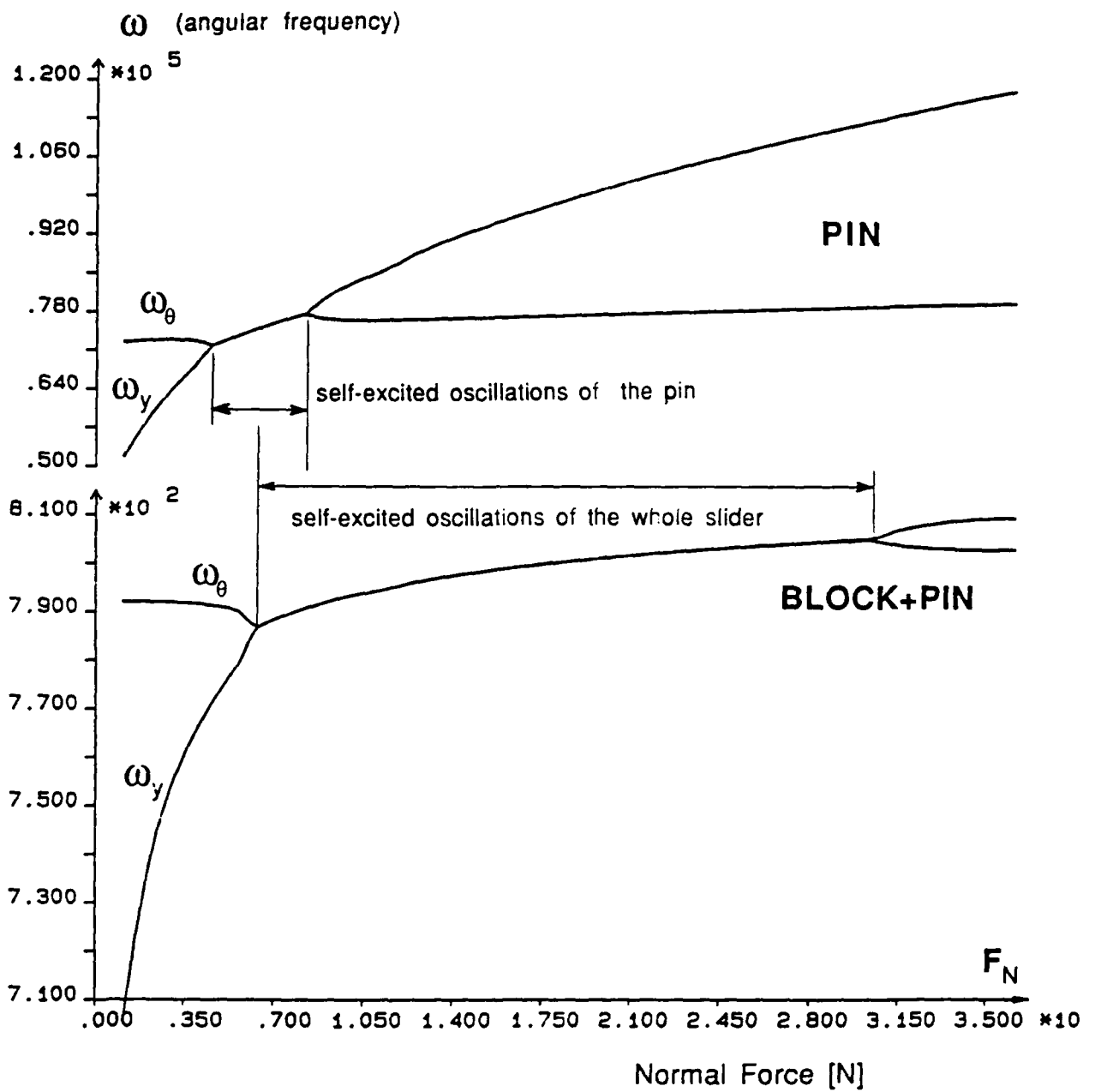


Figure 13: Zone of self-excited oscillations of block and pin at increasing normal load.

It can be seen that self-excited oscillations of pin and block can occur independently and, at certain range of normal loads, there are actually two self-excited modes of different frequencies. For more complex systems this observation can be generalized to predict the whole spectrum of self-excited modes, each of them corresponding to different noises generated by the system.

In the analysis presented so far the disk was assumed to be fixed in the normal direction (due to lack of relevant data). However, in practical systems, shaft and bearings have certain compliance, so that disk can also vibrate during rotation. Therefore the question arises: *Can self-excited oscillations occur due to coupling between motion of the slider and the disk, for example, rotational vibrations of the slider and normal (i. e., along the shaft) oscillations of the disk?* The answer to this question is positive. In order to illustrate possible situations of this type we have perturbed the normal stiffness of the supporting arm (in order to avoid coupling between rotational and normal modes of the slider) and assumed compliant disk support. In particular, for the data:

$$\begin{aligned} K_y^C &= \text{BIG} \\ K_y^B &= 1.37 \cdot 10^7 \text{ kg s}^{-2} \end{aligned} \tag{8.6}$$

we have obtained graphs of frequencies of normal vibrations of the disk and rotational vibrations of the slider presented in Fig. 14 and the zone of self-excited oscillations similar to previously presented examples (but now both disk and slider vibrate).

Again, like in previous cases, this observation can most probably be generalized to more complex multi-component sliding systems and suggests the existence of considerably complex modes of self-excited oscillations, with dynamic coupling between different elements of the system.

The above few examples illustrate sensitivity of the stability of frictional sliding to various parameters of the system. The significance of many of these parameters can be intuitively anticipated by analysis of their influence on frequencies of normal and rotational vibrations. It is of importance to observe that introduction of damping into the system, apart from the overall damping effect, alters frequencies of natural vibrations. Thus, it can, in some cases, actually shift the system out of or *into* the instability zone. The role of damping will be studied in Section 10.

8.3 Transient Analysis of Self-Excited Oscillations

The eigenvalue analysis presented in previous paragraphs estimates dynamic stability of infinitesimal linearized oscillations about the steady-sliding equilibrium position. The closer modeling of the behavior of real systems can be obtained by transient solution of nonlinear equation (5.5) after small perturbation of equilibrium position. Such computations were

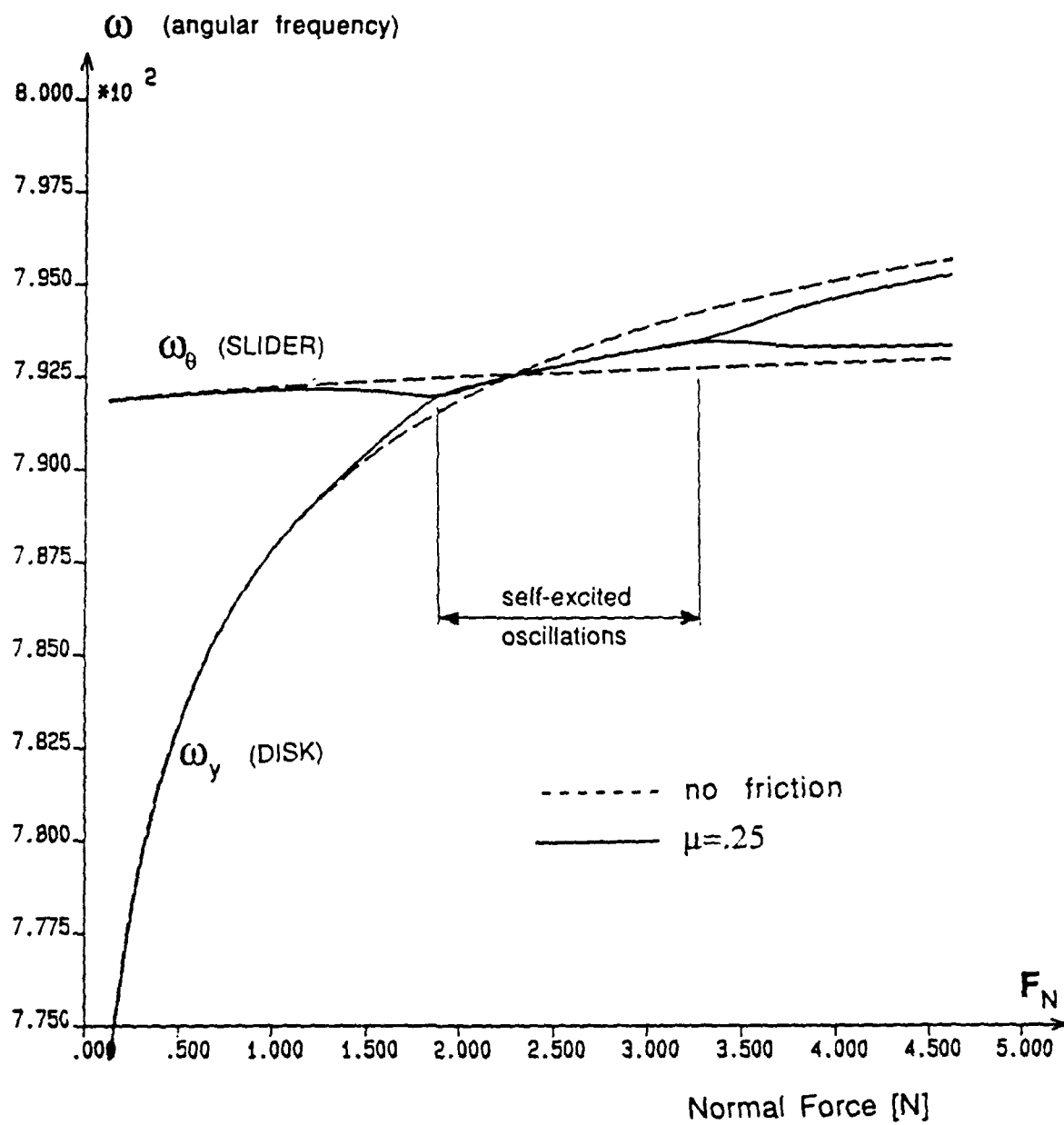


Figure 14: Self-excited oscillations in the case of coupling between normal node of a disk and rotational mode of a pin

performed and confirm predictions of the eigenvalue analysis. For example, Fig. 15 presents a history of normal velocity of the block obtained for our basic set of data presented in Section 8.1 with the following additional parameters:

$$\begin{aligned}
 K_y^{A-C} &= 2.057 \text{ kg s}^{-2} \\
 F_N &= 200 \text{ N} \\
 p_v &= 0.01 \text{ cm/s} \\
 \epsilon &= 0.01 \text{ cm/s} \\
 \Delta t &= 10^{-6} \text{ sec} \\
 \mu_s &= 0.25
 \end{aligned}
 \tag{8.7}$$

In the above data p_v is the small initial perturbation of the normal velocity of the block, ϵ is a regularization parameter in friction law and Δt is a basic time step (automatically adapted during further analysis). It is evident on the figure that oscillations of the system are actually self-excited and grow as a function of time (the same happens with rotational oscillations, not presented here). This behavior complies with the predictions of the eigenvalue analysis presented in Fig. 12. The behavior of the system at fully developed oscillations is analyzed in the next section.

A more detailed analysis, not presented here, suggests that sufficiently large perturbation can, under some circumstances, cause self-excited oscillations of the systems which, according to linearized analysis, are stable. Therefore the actual instability zone may extend slightly beyond the limits estimated by linearized analysis.

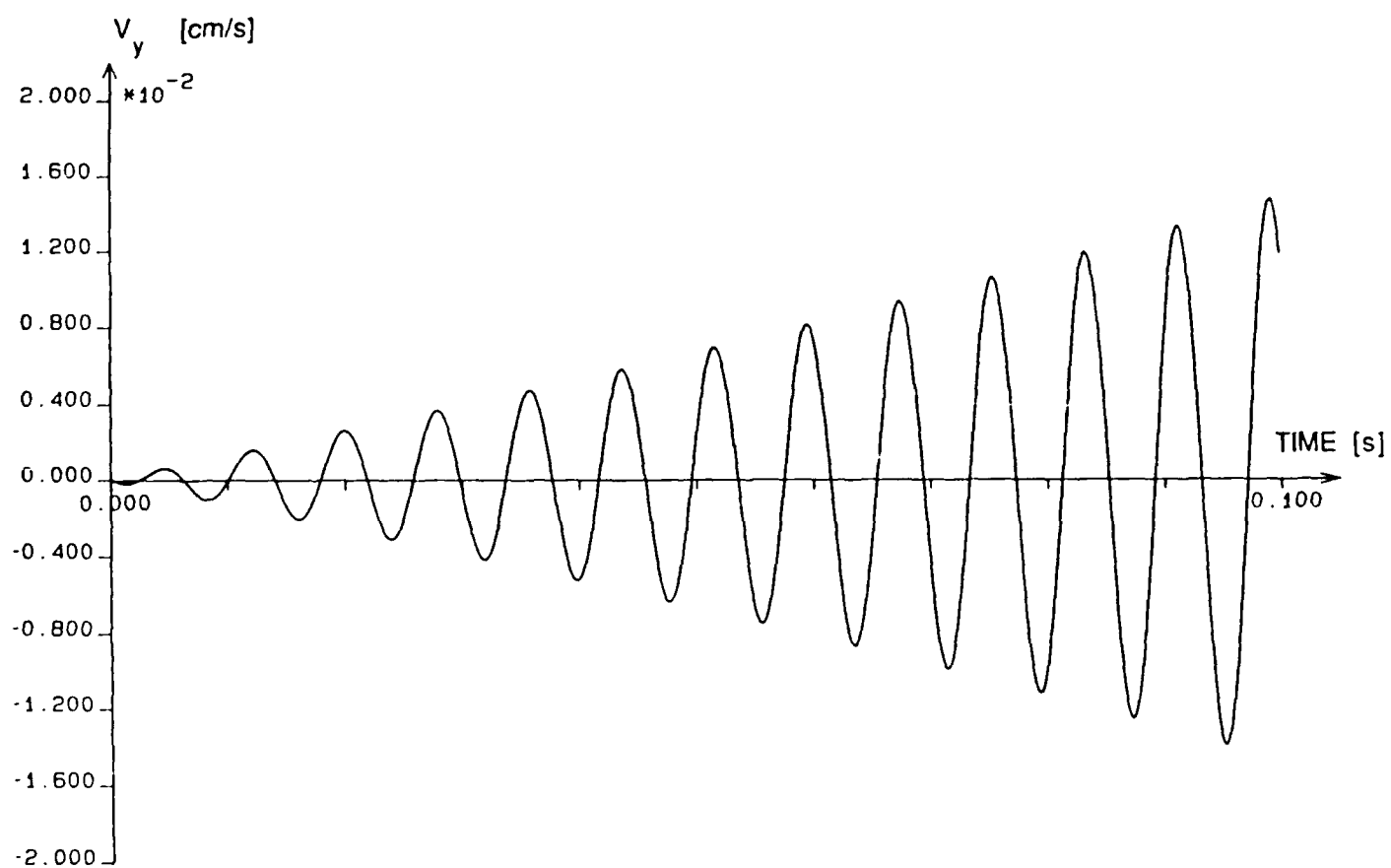


Figure 15: Evolution of normal velocity of the slider at the beginning of unstable vibrations.

9 Introductory Analysis of the Kinetic Coefficient of Friction

In this section a transient analysis of motion of the system after onset of self-excited oscillations will be performed. The particular goal is to analyze changes of the apparent kinetic coefficient of friction in the presence of self-excited oscillations, like those observed in the experiment (Fig. 9, zone IV).

The transient analysis was performed using the Newmark method with an adaptive time stepping procedure presented by Oden and Martins [1985]. Although the Newmark's algorithm is unconditionally stable, the selected basic time step $\Delta t = 5.0 \cdot 10^{-6}$ sec was small enough to model all frequencies of the system, in particular, high-frequency vibrations of the pin: (order of 15 kHz). The motivation was accuracy. Since the reaction from the interface is transferred to the block via the pin, we wanted to follow the motion of the pin with sufficient precision.

Some of the presented computations were also repeated using the Adams-Gears integrator with "arbitrary" user-specified accuracy. The results were virtually identical as those obtained with Newmark's method, yet the Gears integrator was much slower.

As the first example we analyze the motion of the basic model presented in Section 8.1 with the following additional parameters:

$$\begin{aligned}
 K_y^{A-C} &= 2.057 \text{ kg s}^{-2} \\
 F_N &= 200 \text{ N} \\
 \epsilon &= 0.01 \text{ cm/s} \\
 \mu_s &= 0.25 \\
 p_d &= 0.01 \text{ cm} \\
 \Delta t &= 5.0 \times 10^{-6} \text{ sec.}
 \end{aligned}
 \tag{9.1}$$

where p_d is the perturbation of displacements U_y^A, U_y^C from steady-sliding equilibrium position and other parameters were explained before. This set of data, which is referred to as the basic data, corresponds to the zone of self-excited oscillations.

The calculated history of horizontal displacements of the slider (which reflects the apparent friction force) is presented in Fig. 16. It can be seen that the oscillations of the system grow, and after about .3 sec. the friction force temporarily drops. It was verified that this reduction of friction force occurs whenever the sliding of the pin turns into the microscopic

stick-slip motion in phase with rotational vibration of the slider as a whole. However, for this particular model the reduction of friction force is only temporary, which is not in agreement with experimental observations. The main reason is lack of damping in the model which allows for strong high frequency vibrations of the pin itself, which interfere with the stick phase of the microscopic stick-slip motion (the stick contact is disturbed). The behavior of the system, however, changes significantly after the application of small damping into the system, in particular into the pin-to-block connection.

This is illustrated in Fig. 17 which presents horizontal displacements of the block after introducing dashpots into the pin-block connection. The actual damping coefficients were equal to:

$$\begin{aligned} C_x^{A-C} &= 100 \text{ kg s}^{-1} \\ C_y^{A-C} &= 130 \text{ kg s}^{-1} \\ C_\theta^{A-C} &= 100 \text{ kg cm s}^{-1} \end{aligned} \tag{9.2}$$

and give damping close to critical damping for the pin, but only one percent of critical, for the slider as a whole (which consists of the pin and a heavy block). Physically, existence of such a small damping is realistic and well justified.

As can be seen, when the self-excited oscillations reach certain levels, the value of apparent kinetic coefficient of friction drops below the steady-sliding value. Moreover, the motion of the tip of the pin has the form of stick-slip in accord with the rotational vibration of the slider. This can be seen in the next figure, which presents a time history of the relative sliding velocity on the frictional interface.

These results are in perfect qualitative agreement with experimental observations in the zone of self-excited oscillations. However, the actual calculated reduction of the mean value of the kinetic coefficient of friction with respect to steady-state value is about 10 percent, which is much smaller than the experimental results on Fig. 9.

The reasonable explanation of this discrepancy is the fact that the actual value of the coefficient of friction in the self-excited zone is extremely sensitive to the mechanical characteristics of the system. For example, at a stiffness of the supporting arm slightly different than analyzed here, there occurred no reduction of the kinetic coefficient of friction in the self-excited zone (see Fig. 5 in Aronov, et. al [1984]).

In terms of our numerical model it seems that the actual drop of the kinetic coefficient of friction is related to the relative length of the micro-stick phase: the longer the stick, the stronger the reduction of the mean friction force. The duration of stick phase depends in turn on "strength" of the instability, which can be measured by the magnitude of the real parts of eigenvalues of equation (8.4). This fact is confirmed by the example in which

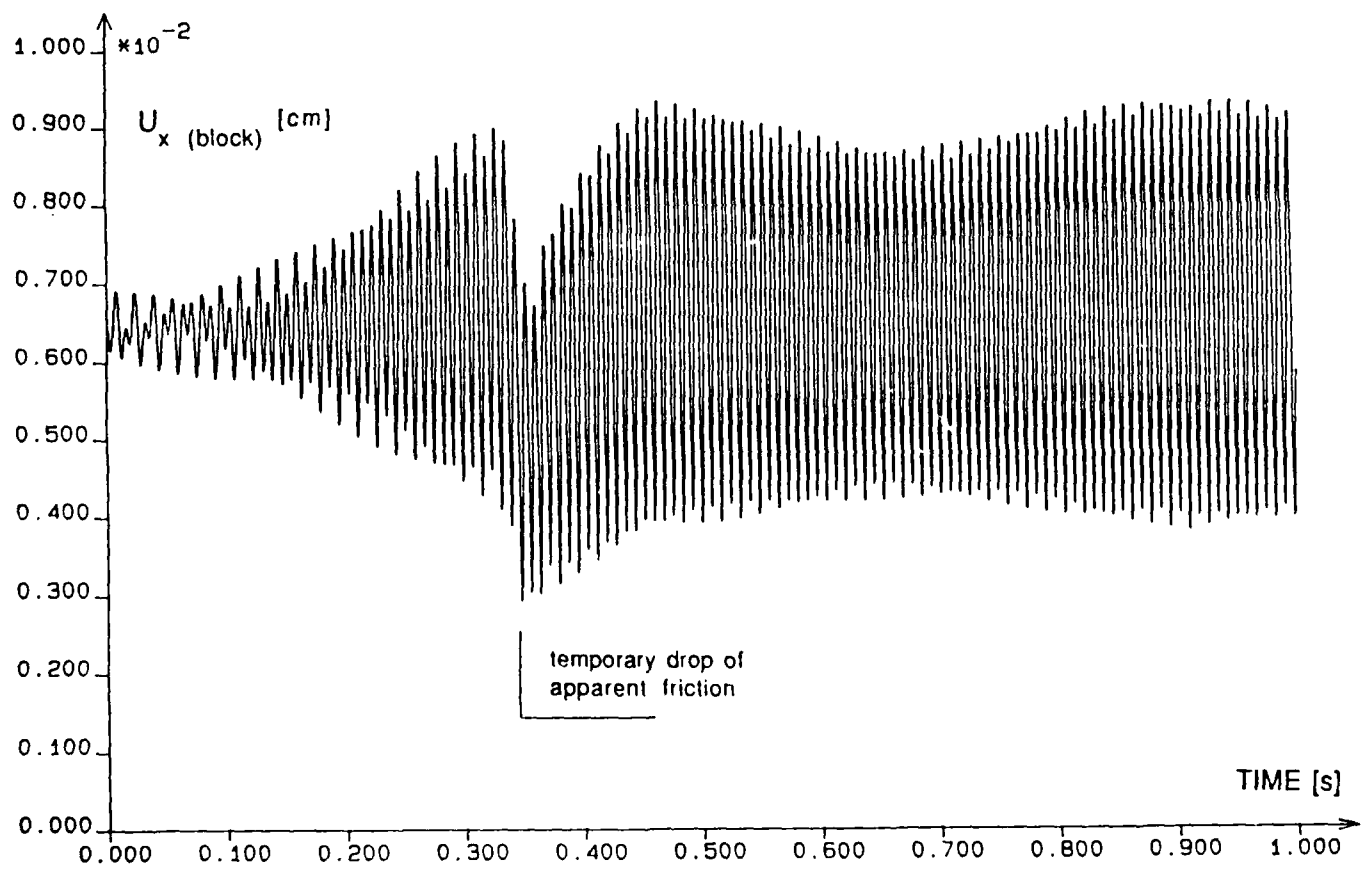


Figure 16: Temporary drop of the friction force in the self-excited zone—no damping.

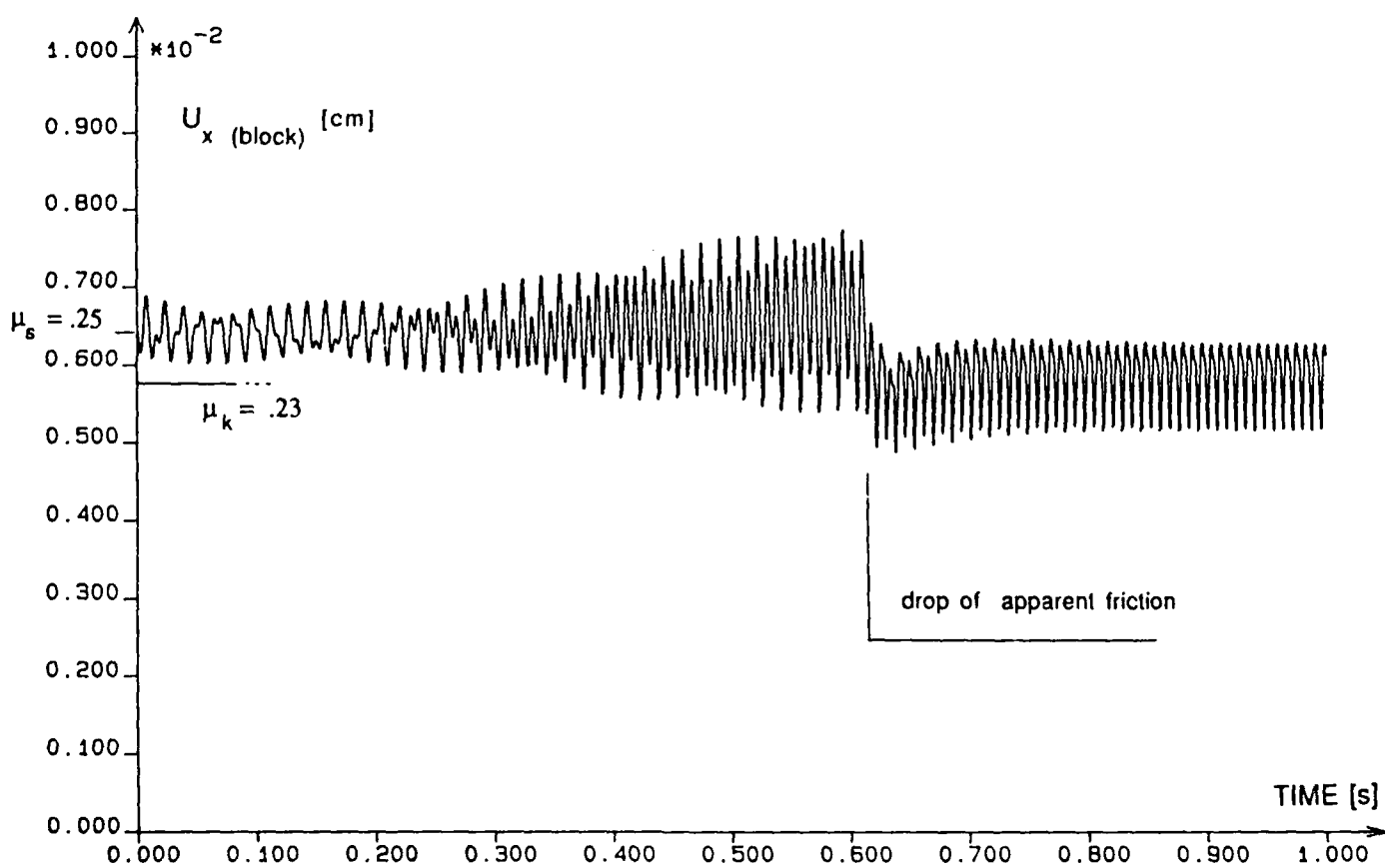


Figure 17: Drop of friction force in the zone of self-excited oscillations.

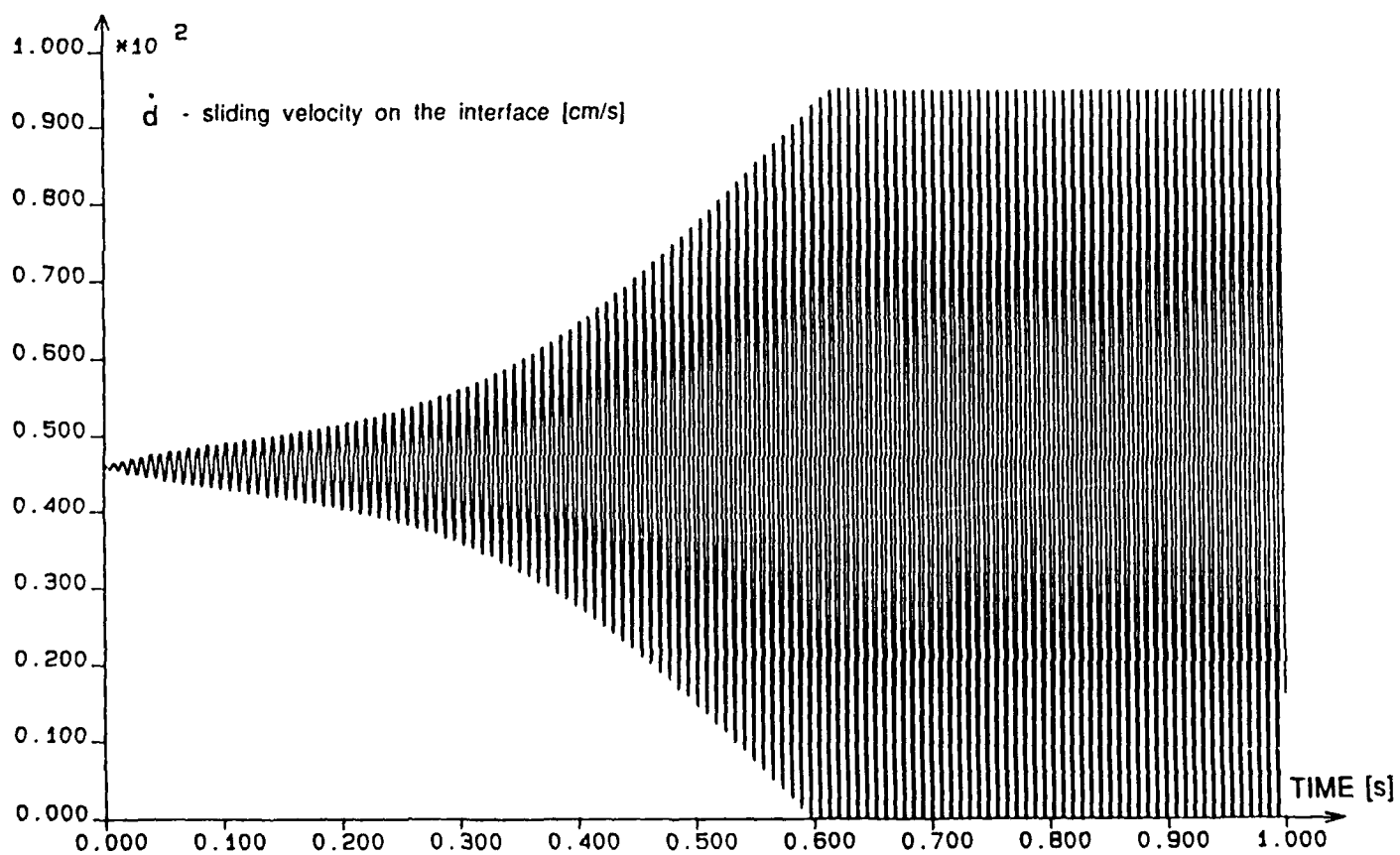


Figure 18: Evolution of Relative horizontal velocity during the self-excited oscillations

the stiffness K_y^{A-C} was changed to the value $2.4 \times 10^6 \text{ kg s}^{-2}$, which made the dynamic instability "stronger" than in our basic example. In this case, the stick phase of microscopic stick-slip motion was relatively longer and, consequently, the reduction of the apparent kinetic coefficient of friction was more significant than in previous examples (compare Figs. 17 and 19).

It is important to observe that, in general, there seem to be two basic reasons which limit the growth of oscillations in the self-excited zone. Since the instability is due to coupling between normal and rotational modes of the system, the growth of oscillations can be limited by some mechanisms, which alter one of the above frequencies. These mechanisms can be in particular:

- occurrences of stick-slip motion of the tip of the pin, which tends to alter rotational frequency of the slider (like in examples considered here),
- occurrence of normal jumps of the slider, which tends to alter its normal frequency.

Depending on the mechanical characteristics of the system and sliding velocity one of these mechanisms is actually activated. If it happens to be normal jumps of a slider, there will be no microscopic stick-slip and, consequently, no perceptible drop of the apparent coefficient of friction (like in Fig. 5 in Aronov, et. al [1984]).

In the analysis presented so far the model of the interface was relatively simple, with no interface damping introduced and the coefficient of friction independent of the normal load. As it was mentioned in Section 2, these phenomena can, in the presence of vibrations, introduce additional changes to the coefficient of friction.

Unfortunately, for the system under consideration the load dependence of static coefficient of friction or amount of damping on the interface were not known. However, for illustrative purposes, we have solved a few examples, with necessary coefficients based on other sources.

In the first example we introduce some damping on the interface by specifying in the interface constitutive law (3.31) $b_{N-} = 0, b_{N+} = 3.0 \times 10^{12} \text{ kg cm}^{-4} \text{ s}^{-3}$ with $l_n = 2.5$. The values of coefficient b_N are estimated according to remarks presented by Tworzydło and Becker [1988]. At the relative normal velocities on the interface attained during self-excited oscillation (order of 1.0 cm/s) they produce damping force of the order of 10 percent of the elastic response of the interface. The plot of the calculated history of tangential displacements of the slider (which reflect the measured apparent friction force) is presented in the Fig. 20. It is seen that as the oscillations grow, the apparent coefficient of friction decreases and the final reduction of friction is more significant than in the case without interface damping (15 percent compared to 11 percent).

Additional drop of friction force was also obtained when the coefficient of friction was assumed to be dependent on normal load. In particular, it was assumed that at load 200 N the coefficient of friction expressed in equation (3.35) is equal to .25, while at loads decreasing

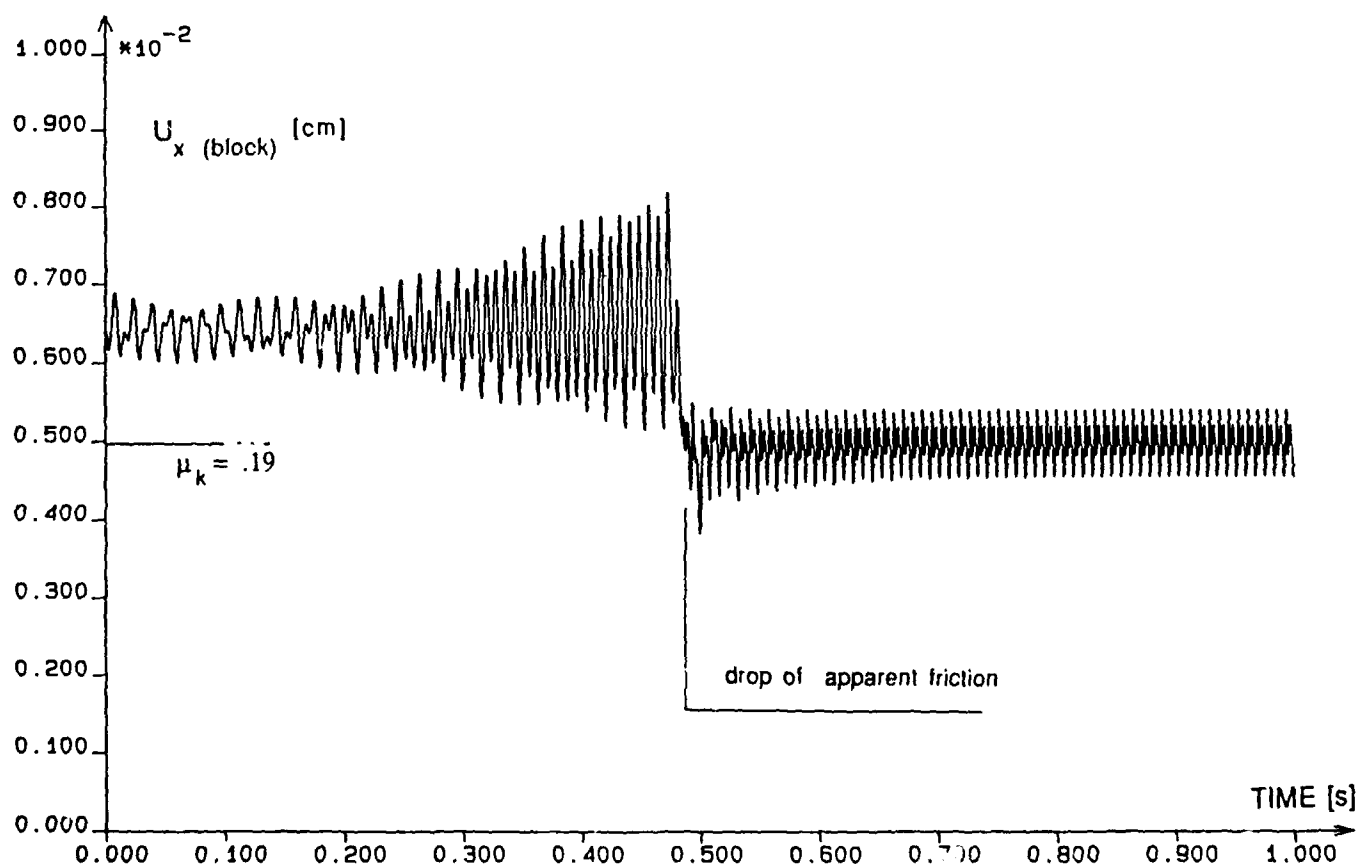


Figure 19: Drop of friction force due to self-excited oscillations—"stronger" instability than for basic data set.

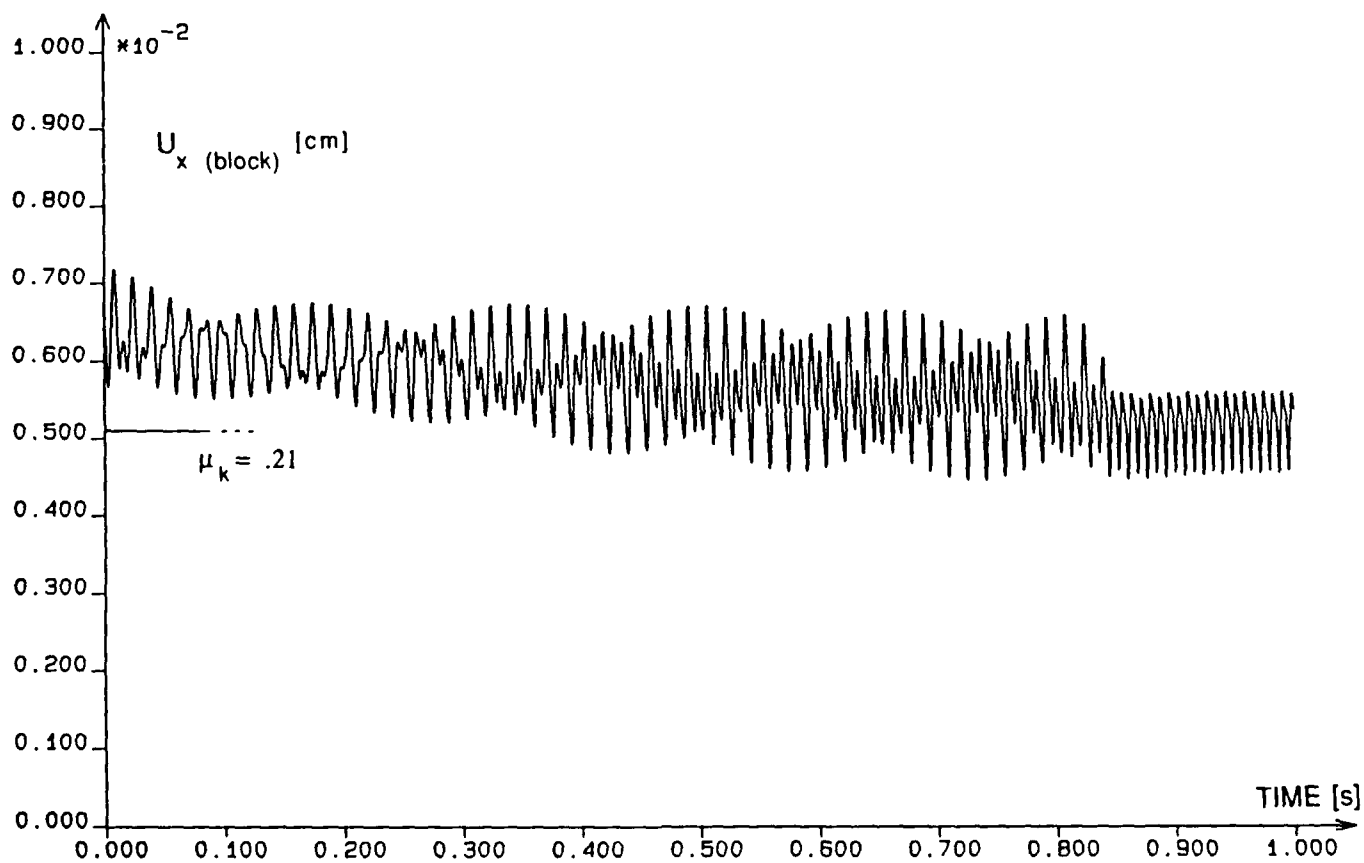


Figure 20: Drop of the friction force in the self-excited zone—nonsymmetric damping on the interface

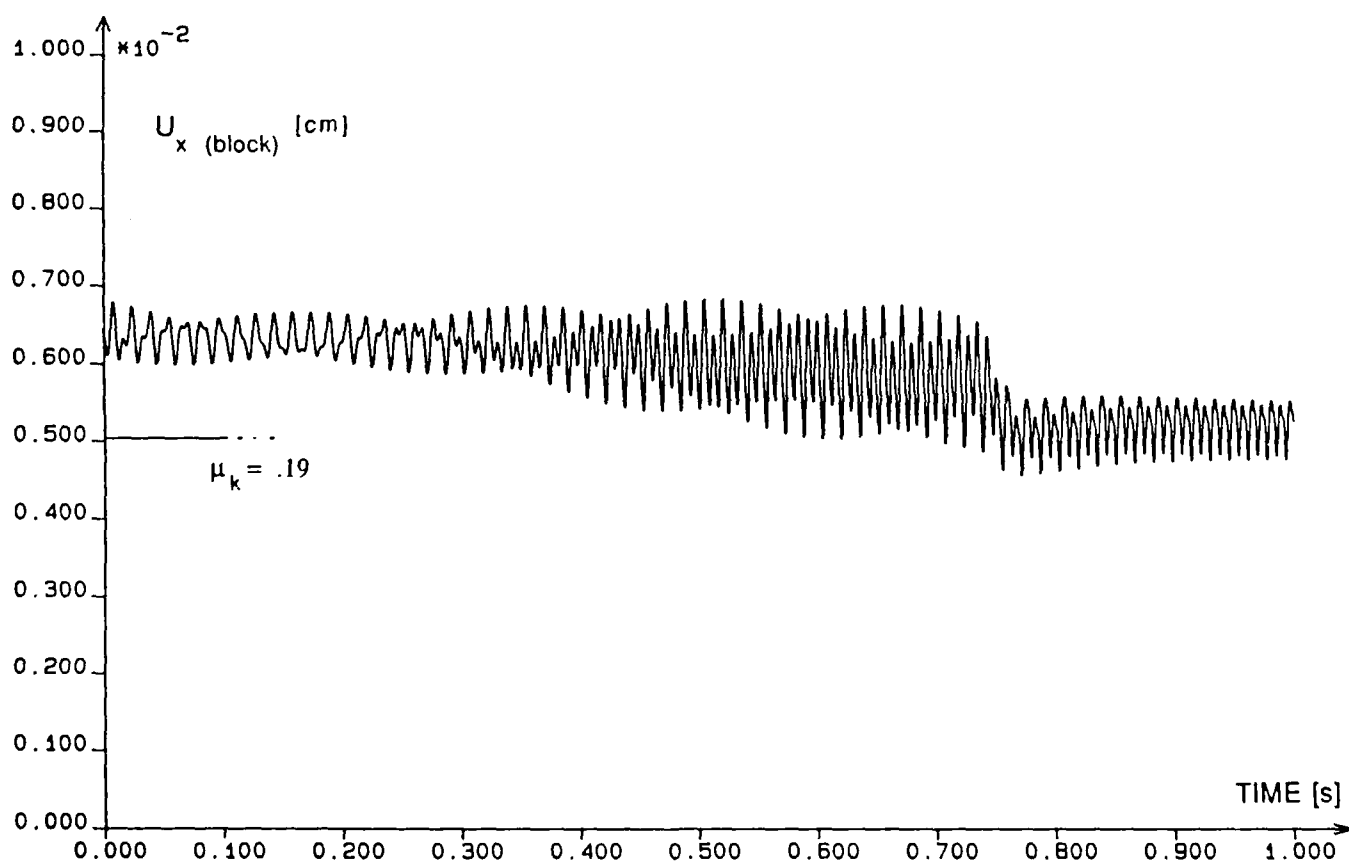


Figure 21: Drop of the friction force in the self-excited zone—load dependent coefficient of friction

to 0 N it linearly approaches 0.23. For such defined data, the time history of friction force is presented in Fig. 21. Again, the drop of coefficient of friction is more significant than in our basic case presented in Fig. 17, namely 21 percent compared to 11 percent.

The above examples show that various parameters of the interface and the system as a whole affect the final value of the coefficient of friction in the presence of self-excited oscillations. Hopefully, exact estimates of all involved parameters, which were impossible in this study, should provide even better agreement between numerical and experimental results. This issue requires further, more detailed study.

It should also be noted that the above analysis was focused on the modeling of coefficient of friction only in the unstable zone, where the coupling between normal and rotational vibrations was the dominant source of vibrations.

Analysis of the kinetic coefficient of friction in steady-sliding would require proper modeling of vibrations due to interlocking of surface asperities, which provides excitation of vibrations in this zone.

10 Influence of the Angle of Attack on the Stability of Frictional Sliding

In the preceding sections we have analyzed behavior of the system under the assumption that the angle of attack was equal to zero, which is equivalent to saying that the surface of the pin is parallel to the surface of the disk prior to the application of loads.

In the experiment this condition was assured by bringing the pin into contact with the revolving disk at small normal load and allowing it to run for a distance of approximately 12–15 km. The angles of attack different from zero were not analyzed experimentally.

However, in our numerical modeling we have also considered various angles of attack, ranging from -10° to $+10^\circ$, where positive angle corresponds to clockwise rotation of the slider in Fig. 7 (in actual computation the initial rotation was actually imposed on the disk). These angles are within the range of applied theory, in which it was assumed that $\sin \theta \approx \theta$ and $\cos \theta \approx 1$. At various angles of attack the described eigenvalue analysis before was used to calculate forces corresponding to the onset of the self-excited oscillation and thus the stability and instability zones were defined for angles of attack ranging from -10° to $+10^\circ$ and normal forces up to 500 N .

The results obtained are presented in Fig. 22, which also shows schematically the configuration of the slider in each of the presented zones. It can be easily observed that generally the self-excited oscillations occur when the point of contact is ahead of the center of the mass of the slider, so that the slider "stumbles". More precisely, presence of friction on the interface is the reason for counterclockwise rotation of the slider, which, in such a configuration of the slider, leads to growth of normal force due to inertia of the slider. This in

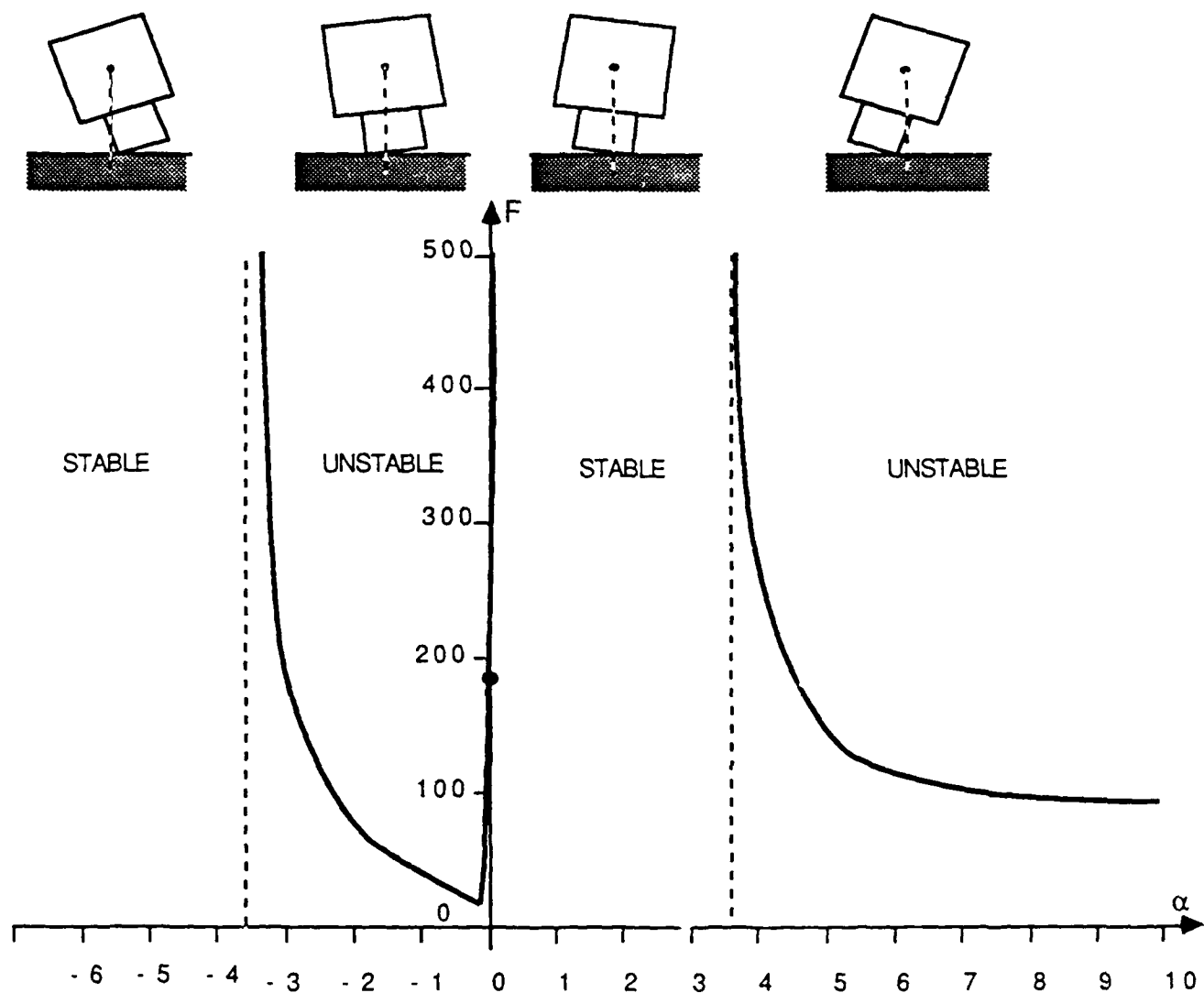


Figure 22: Zones of self-excited oscillations at various angles of attack

turn increases friction force and causes further rotation. Such a coupling between rotational normal vibrations is the reason for high-amplitude self-excited oscillations.

On the other hand, if the point of contact is behind the center of the mass of the slider, the sliding is steady. The reason is that in this situation rotation of the slider, due to friction, causes decrease of normal and friction forces, so there is no coupling between normal and rotational modes.

An especially interesting zone is in the vicinity of the zero angle of attack. It can be seen that if this angle is exactly zero, the onset of self-excited oscillations corresponds to normal force 185 N (just like in Section 9). However, this value changes drastically when the slider is rotated, so that at positive angles of attack greater than 0.005° , the sliding is steady at all analyzed loads.

This sensitivity is somewhat stronger than one would expect in this case (although there exists no corresponding experimental evidence). The possible explanation can be in the fact that our computational model was perfectly "precise", while real systems have various imperfections, which somewhat "relax" the sharpness of response of the system.

Moreover, the presented analysis is based on a simplified estimate of stability of infinitesimal rotations about the equilibrium position. As stated before, if the fully nonlinear system is analyzed, the instability zone may extend somewhat beyond boundaries established via eigenvalue analysis.

In order to verify this we have performed a full transient analysis for some selected values of angle of attack in the vicinity of $\alpha = 0$. This analysis revealed that the instability zone near $\alpha = 0$ actually reaches values of angle of attack up to 0.05° (one order higher than from the eigenvalue analysis). This value, although still quite small, appeals better to our intuition than the previous result.

The study presented in this section confirms well known sensitivity of stability of frictional sliding to the angle of attack, which is one of the reasons of poor reproductivity of the results of various frictional experiments (in the experiment of Aronov, et al. [1984], this reproductivity was assured by extensive pre-sliding prior to actual measurements). For the future, more precise comparisons with relevant experiments are required.

11 Conclusions

In this work numerical modeling of frictional behavior was presented, particularly stability of mechanical sliding systems.

The good qualitative and quantitative agreement of numerical and experimental results confirms the importance of mechanical characteristics of a system in frictional sliding and applicability of the presented numerical approach.

The identified principal parameters affecting the occurrence of self excited oscillations,

namely frequencies of normal and rotational vibrations, provide an intuitive insight into the mechanism of self-excited vibrations and guidance in the way of avoiding them in the design of sliding systems.

Although the analyzed model was relatively simple, the observations and methods used in this work can be applied to analysis of more complex systems, possibly discretized by the finite element method. However, as happened even in this simple model, the coupling between normal and rotational vibrations can, in general, take the form of a variety of different self-excited modes occurring in one sliding system.

In particular, for contacting elastic bodies, the self-excited oscillations can have a localized form, with only certain parts of contacting bodies actually vibrating. This would require careful analysis of all possible modes in the system.

The presented analysis reveals extreme sensitivity of the possible onset of self-excited oscillations to all parameters of the system. For example, if the compliance of the force transducer sandwiched between pin and block is disregarded, the self-excited oscillations occur at completely different ranges of loads than in the experiment.

It should be pointed out that the model of the interface (generalization of Oden/Martins model) was relatively simple, with purely elastic normal response and linear damping term representing all dissipation effects on the interface. In the future analysis, the more general viscoplastic model of interface should be applied (possibly based on micromechanics analysis). In addition, such models should be combined with appropriate descriptions of lubricants, which reduce the coefficient of friction on the interface and introduce considerable amounts of damping.

The presented introductory analysis of the kinetic coefficient of friction in the self-excited zone confirms the fact that the difference between static and kinetic friction is not an intrinsic property of the interface, but also depends on the overall nature of the motion of the systems, in particular normal vibrations, which accompany virtually every sliding contact. Moreover, in the presence of normal vibrations, various phenomena, summarized in Table 1, influence the final value of the kinetic coefficient of friction.

It is of importance to note that the transient analysis of motion of the slider is extremely expensive, since the macroscopic motion of the slider (observed in the scale of seconds) results from a complicated nature of high frequency vibrations of all components of the system. Thus, hundreds of thousands of time steps are required in order to model even short periods of motion of the slider. Therefore, for models with higher numbers of degrees of freedom (like Finite Element Models), a more efficient approach to the time integration should be devised or enormous computation times can be expected.

In summary, in our opinion, further progress in modeling of dynamic friction phenomena will require:

- evaluation of more adequate phenomenological models of the frictional interfaces

(based on a micro-asperity analysis and consistent with experimental results),

- further development of numerical techniques used in modeling of the motion of the sliding systems,
- development of a solid mathematical basis for problems considered and numerical methods applied,
- further comparison of numerical results with adequately designed experiments.

CHAPTER 4

NEW MODELS OF CONTACT AND FRICTION OF ROUGH SURFACES

1. INTRODUCTION

In this paper, we develop a new theory of contact and friction of dry rough surfaces. The theory involves an asperity-based model of metallic interfaces in which surface topography is characterized by a random distribution of asperity heights and curvatures. The surface roughness is assumed to be isotropic, so that a Gaussian distribution of asperity heights can be assumed, and each asperity is modeled as a paraboloid with quite general mechanical properties. Elastic, viscoplastic and elastoplastic response of the asperities, the latter with strain-hardening and load reversal, can be accommodated in the model. The effects of adhesion through inter-molecular forces is also taken into account. Sliding resistance is modeled by a micro-mechanical constitutive equation for shear stress that is suggested by a reasonably large volume of experimental results. The final macromodel of friction and contact is deterministic and is characterized by constitutive equations for normal and shear stresses on the material interface. The theory developed here represents a substantial generalization of theories of contact and friction reported over the last two decades.

As is well known, attempts at developing mechanical theories for the sliding resistance of contacting bodies date back nearly three centuries to the primitive work of AMONTONS [1699], and some historians suggest that daVinci proposed models of friction over a century earlier. Classical works of COULOMB [1781] pointed to the existence of asperities, perturbances on the contact surface, which were responsible for some resistive forces. He introduced the notion of static and dynamic coefficients of friction. This work, and the pioneering work on elastic contact by HERTZ [1885] had a profound impact on research on contact and friction over the last century.

The classical papers of AMONTONS and COULOMB, of course, necessarily dealt with global descriptions of friction on effectively rigid bodies, since an acceptable notion of stress and deformation were not to be put forth until the early part of the nineteenth century. Also, we note that HERTZ did not consider friction effects but only with the mathematical description of the problem of indentation of one smooth elastic body into another. Nevertheless, the HERTZ solution for elastic contact has been used as a basis for numerous models of elastic asperities designed to simulate the deformation of material interfaces.

In a recent paper, TABOR [1981] presented an excellent summary of the current knowledge of friction effects on surfaces and pointed to some of the principal mechanics responsible for friction on dry rough surfaces. TABOR pointed to three primary factors contributing to the nature

of frictional resistance: 1) the true contact area, as opposed to the nominal or apparent contact area, adhesive forces due to inter-molecular attractions, and 3) sliding resistance, which can be determined by a number of factors, but may in particular depend upon the shear strength of interfacial junctions or films. Many mathematical models of contact and friction have attempted to simulate one or more of these characteristics.

The first contact model that was constructed to predict the true contact area can be found in a paper by ABBOTT and FIRESTONE [1933], in which the contact surface was simulated in a network of spheres that are truncated upon indentation into a hard flat. By knowing the hardness of the softer of the two materials in contact, an estimate of the true contact area could be made, assuming perfectly plastic deformations. Similar results, plus a large volume of experimental evidence together with the notion of interface junctions of asperities and the importance of the shear strength of interface materials on the magnitude of the coefficient of friction, were discussed in the two volume treatise by BOWDEN and TABOR [1950, 1964]. These works, which summarized a great deal of experimental and empirical data on friction that existed up to the early 1960's, were important predecessors to modern attempts at developing continuum theories of deformable interfaces and of friction.

An important advance in development of asperity-based models of contact is represented by the pioneering paper of GREENWOOD and WILLIAMSON [1966], in which the rough surfaces were viewed as a randomly distributed population of elastic asperities with randomly distributed asperity heights. Each asperity was assumed to be spherical and elastic and its deformation properties governed by the Hertz solution for elastic contacts. Experimental evidence was provided to support the assertion, now widely held by tribologists, that for normally isotropic engineering surfaces, a Gaussian distribution of asperities heights generally exists. In such models, there are no micro-frictional effects on the asperities, such effects leading to second-order changes in contact pressure, a result established nearly two decades earlier by MINDLIN [1949]. In a related paper, GREENWOOD and TRIPP [1967] showed that contact of two rough surfaces with Gaussian distributions of asperity heights on which asperity contacts were misaligned was equivalent to a single elastic surface with a Gaussian distribution of asperity heights impinging on a rigid flat. The use of such statistical representations of surface topography has since become a popular approach in modeling both elastic and inelastic contacts.

The GREENWOOD-WILLIAMSON model was based on the assumption that only the asperity height was a random variable, and that the radius R of each peak was constant. Several generalizations of such random topography models appear in the literature of the 1970's. The paper of WHITEHOUSE and ARCHARD [1976] extends the random-asperity models to include random heights and curvatures, and NAYAK [1971] provided a general approach to random surface

modeling using notions of joint probability distribution functions. In this same vein, we mention the work of BUSH, GIBSON, and THOMAS [1975], who derived a joint probability distribution density function for random asperity heights and curvatures of a random population of elliptic paraboloids in elastic contact with a smooth rigid flat.

Such random-microtopography models that employ a deterministic function for asperity peak shapes are called *asperity models*. One source of possible inconsistency in such models has to do with the fact that a Gaussian distribution of asperity heights and curvatures for a given asperity shape may lead to a non-Gaussian cumulative probability distribution of the surface height, an unrealistic result for most "engineering surfaces." This problem was addressed by HISAKADO [1974] and HISAKADO and TSUKIZOE [1974], by assuming a Gaussian PDF (probability density function) for surface heights, with a given deterministic asperity shape, and then deriving the PDF for peak heights. HISAKADO [1974] assumed a paraboloidal asperity shape and HISAKADO and TSUKIZOE [1974] a conical shape. FRANCIS [1977] points out that the HISAKADO models may lead to unrealistic PDF's for asperity heights, since they may be strongly dependent on the asperity shape and may become negative for paraboloidal and conical shapes.

Extensions of asperity-based models to microcontact deformation laws involving elastoplastic deformations were first contributed by HISAKADO [1974] and HALLING and NURI [1975] account for plastic deformation of the interface by assuming that a rough surface deforms elastically while contacting a nonlinearly elastic flat, representing strain-hardening, with each microcontact defined by a fully-plastic spherical indentation. Significant generalizations of these types of asperity models can be found in the detailed studies of FRANCIS [1976,1977], who introduces the notion of the *sum surface*, discussed later in the present paper. This enables one to model Gaussian engineering surfaces with asperity shapes that are paraboloidal only at their vertices, but which have random heights and curvatures, using the joint PDF of NAYAK [1971]. Moreover, FRANCIS [1977] also takes into account elastic and fully plastic deformations, with strain-hardening, using functions determined empirically from spherical indentations of various metals.

We also mention that an extension of the GREENWOOD-WILLIAMSON model of spherical asperities with Hertzian elastic contacts, constant radii, and random heights to cases in which a transition to perfectly plastic deformations occur was recently proposed by CHANG, ETSION, and BOGY [1986a].

As pointed out by TABOR [1981], adhesion effects can be very important in dry contacts of highly polished metallic surfaces. Several models of adhesion have been proposed in the literature, and we mention the models of JOHNSON, KENDALL, and ROBERTS [1972], (the JKR model), and DERJAGUIN, MULLER, and TOPOROV [1975, 1980, 1983], (the DMT model). Experiments have indicated that the JKL model seems to provide better agreement with experiments

on adhesion for polymers and, for example, rubber contacts, whereas the DMT (particularly their extended model [1983]) produces results in better agreement with experiments on metals.

Various models of shear resistance in films and lubricants can be found in the paper of BRISCOE, SCRUTON, and WILLIS [1973], BRISCOE and TABOR [1975], and others; see also FRANCIS [1977]. Recently, CHANG, ETSION, and BOGY [1986a,b,c] extended the GREENWOOD-WILLIAMSON model to include plastic deformations accounted for adhesion in a contact model, using the DMT theory, and assumed that impending sliding would occur when the largest principal stress at a metallic junction reached value determined by the Von-Mises yield criterion. These models assumed a spherical asperity shape with constant radius and no strain-hardening of the material upon plastic deformation of asperities.

We note that most of the references cited above dealt with attempts to model either contact without sliding motion, or purely static or quasi-static friction effects. There are, however, many dynamic friction phenomena that can be adequately modeled by continuum models of interfaces that can adequately simulate the normal compliance of material interfaces. The importance of compliance of an interface normal to the direction of sliding in various dynamic friction effects, such as friction damping, and stick-slip-motion was suggested in the experiments of TOLSTOI [1967] and was instrumental in the recent work of ODEN and MARTINS [1985] on this subject. ODEN and MARTINS showed that if a constitutive equation for the interface, characterizing elastic normal contact, was introduced which was consistent with early experimental results and with asperity-based models such as that of GREENWOOD and WILLIAMSON [1966], then a tractable mathematical theory of frictional contact could be obtained, which could produce results in good agreement with experimental data. The ODEN-MARTINS model considered only elastic contacts and, therefore, could not account for changes in the frictional characteristics of an interface due to yielding or failure of asperities.

On the mathematical side, the first attempts at formulating the problem of unilateral contact in elastostatics is attributed to SIGNORINI [1933, 1950], although the mathematical theory of unilateral contact in elasticity was not put together until nearly three decades later in the work of FICHERA [1966]. FICHERA's work showed that the proper mathematical framework for unilateral contact problems in elasticity, and in solid mechanics in general, lay in the theory of *variational inequalities*. His work led to numerous studies of the theory of variational inequalities during the 1960's and 1970's. The first attempt at a mathematical analysis of the Signorini problem with friction appears in the important memoir of DUVAUT and LIONS [1972], where frictional resistance is introduced pointwise on the contact surface using Coulomb's law. These authors point out that use of this classical constitutive equation does not lead to a tractable theory. Indeed, DUVAUT and LIONS announced that the elasto-dynamics problem with Coulomb friction was an

open problem in the theory of variational inequalities and partial differential equations.

Some progress toward resolving this open problem was made by NECAS, JARUSEK, and HASLINGER [1980], and then only for the static case and under special assumptions. One problem with the use of the Coulomb model, aside from its poor correlation with physical experiments, is that it does not lead to an unambiguous characterization of stress on the contact surface; see DUVAUT [1983]. This led ODEN and PIRES [1982, 1983] and DEMKOWICZ and ODEN [1982] to introduce non-local models of friction which, from the mathematical point of view, exhibited the compactness requirements needed for a tractable mathematical theory (in the sense that it was then possible to resolve the issues of existence of solution and uniqueness of solution with the small friction, for contact problems in elastostatics). However, the simple nonlinear constitutive models of ODEN and MARTINS [1985a, 1986] and MARTINS [1986] also provided an appropriate mathematical framework for such theories while, at the same time, provided a theory consistent with known properties of elastic contacts which yielded predictions of macro-response in general agreement with available experimental results. Recently, the mathematical analyses of extensions of the Signorini problem to cases with nonlinear interface laws and friction have been contributed by RABIER, MARTINS, ODEN, and CAMPOS [1986] and to elastodynamics problems with friction by MARTINS and ODEN [1987]. Further references on contact problems and their approximation can be found in the book of KIKUCHI and ODEN [1987].

In the present paper, we develop an asperity-based model of rough, isotropic, engineering surfaces with random microtopography for which the mechanical response of the asperities may be elastic, viscoelastic, rate-dependent (visco-) plastic, or rate-independent elastoplastic, with strain-hardening and load reversal. The effects of adhesion due to inter-molecular attraction are also included, as is sliding resistance due to either the fracture of interface junctions or to shear stresses on interfacial films. The microtopography model makes use of the notion of the sum surface, following FRANCIS [1977], and adhesion is described using the ideas of DERJAGUIN, MULLER, and TOPOROV [1975, 1983] (the DMT model) and CHANG, ETSION, and BOGY [1986b].

The phenomenological constitutive laws for viscoplasticity put forth by BODNER and PARTON [1972] are used to model the mechanical response of the interface material. For situations in which sliding resistance is governed by shear resistance of thin lubricant films, microcontacts are characterized by constitutive equations suggested by the experiments of BRISCOE, SCRUTON, and WILLIS [1973] and BRISCOE and TABOR [1975], as discussed by FRANCIS [1977].

The analysis of a typical asperity is, in itself, a difficult problem in contact mechanics in that

it involves the unilateral contact of a viscoplastic material body subjected to inter-molecular forces which act on surfaces that change in measure during the deformation. We use a new finite element method for the numerical solution of these nonlinear evolution problems and compute time histories of the micro-contact areas and net forces as functions of time for specified macro-contact load histories. These are then fed into the probabilistic microtopography models to yield, as expected values, deterministic macro-contact constitutive equations for the interface.

The resulting contact and friction model generalizes all asperity-based models previously proposed in literature. It reduces, under appropriate assumptions, to the models of GREENWOOD and WILLIAMSON [1966], FRANCIS [1977], CHANG, ETSION, and BOGY [1986a,1986b], and it can yield the interface constitutive laws of ODEN and MARTINS [1985]. Moreover, important special cases of the theory are known to characterize mathematically tractable problems (well-posed problems in a broad sense) in static and dynamic contact mechanics, and to be capable of yielding results consistent with experimental evidence on contact and friction of engineering surfaces. Finally, the theory can be used to model changes in the contact and friction characteristics of surfaces brought about by yielding and viscoplastic deformations.

2. SURFACE TOPOGRAPHY MODELS

We begin by considering the contact of two deformable bodies, I and II, over a nominal contact area A_0 , as illustrated in Fig. 2.1. An element of unit nominal contact area is isolated for study, as indicated in the figure. The average stress vector Σ over the unit contact area has components of force of magnitude P and Q normal and tangential to the unit area, respectively. The situation is equivalent to that of two typical coupons of surface material, one taken from the material near the contact surface of each body, pressed together with a force P normal to the tangent plane at the center of the coupon interface and simultaneously subjected to a shear force Q tangent to the plane. The bulk deformations of bodies I and II are ignored, our aim being only to characterize the mechanical properties of the contact interface. The nominal unit surfaces in contact are, for the present, assumed to be initially flat and parallel to one another.

2.1 Random Microtopography Model. It is standard practice to depict the approximate profile of rough engineering surfaces with a profilometer or stylus, drawn across the surface, which generally yields a jagged profile with an exaggerated vertical scale of the type shown in Fig. 2.2 a.

We consider two such opposing surfaces 1 and 2 which are to ultimately come in contact. Reference planes defining the mean asperity height of each surface profile are established, and we

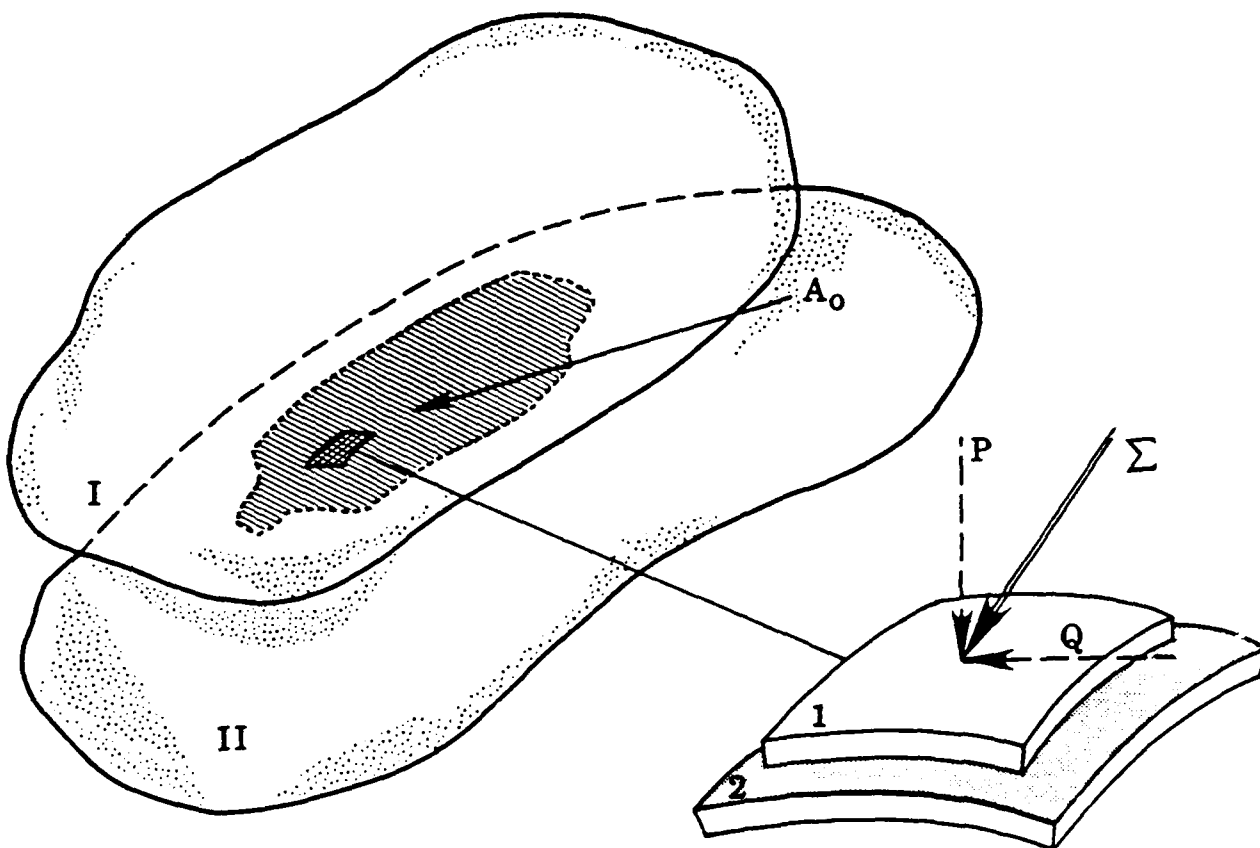


Fig. 2.1 Contacting bodies and coupons near the contact interface subjected to an average stress vector of magnitude Σ .

characterize the shape of each profile by introducing functions z_1 and z_2 , given the height of asperities above the respective reference planes, i.e. the functions $z_i = z_i(x,y)$, $i = 1,2$, with (x,y) a point in the parallel mean-height reference planes, define the profiles of the rough material surfaces 1 and 2, respectively. The distance h between plans is the separation of the surfaces, and the distance between actual opposing material points is denoted s . Thus, as a point (x,y) on the reference plane, we have

$$s = h - z \quad (2.1)$$

where z is the *sum surface* (Cf Francis [1977]),

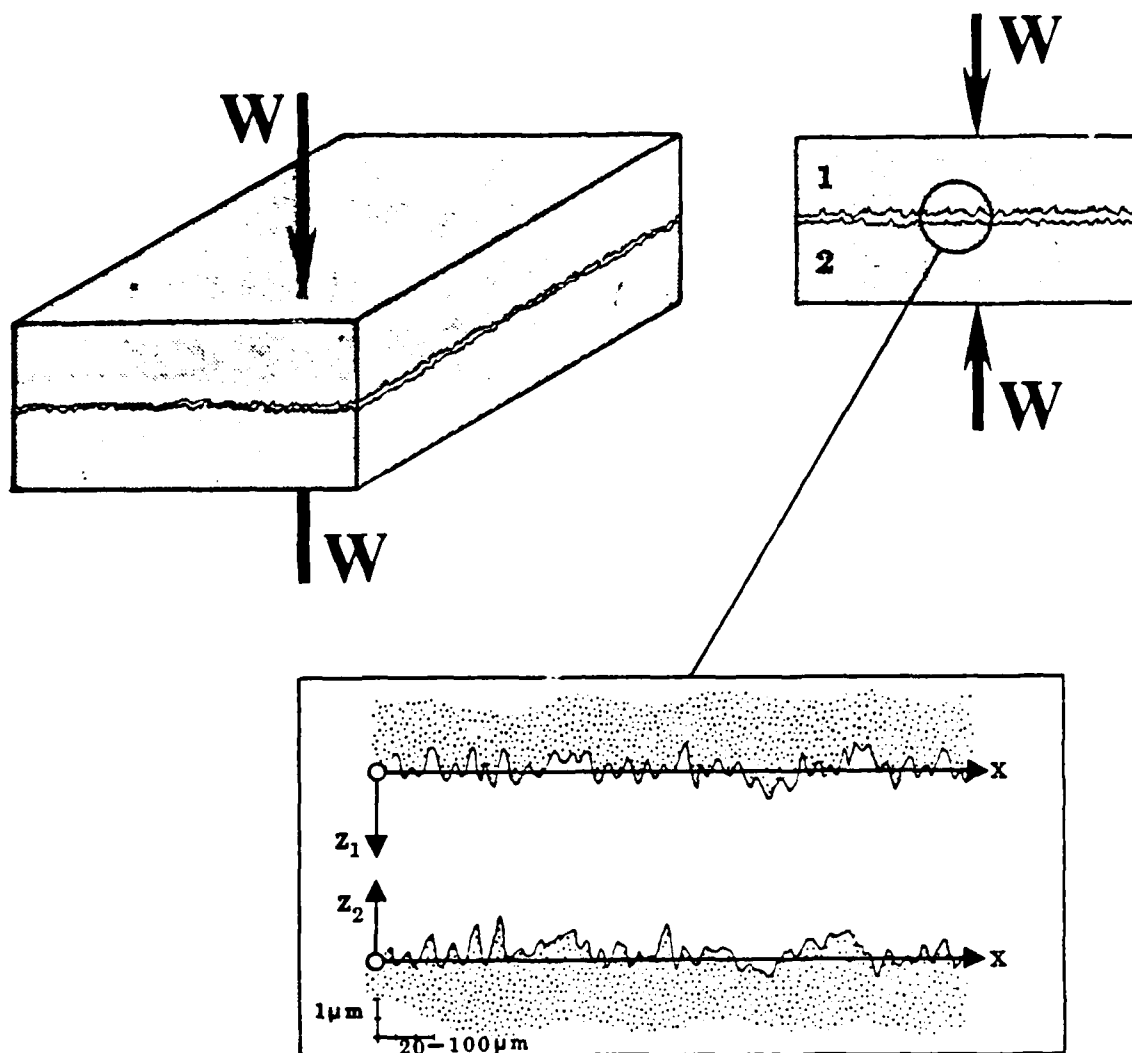
$$z = z_1 + z_2$$

FRANCIS has pointed out that, from the simple fact that the sum z of the surface heights appears in the geometric relation (2.1), the situation is equivalent to that of a single deformable surface of height $z = z_1 + z_2$ approaching a rigid flat, as suggested in Fig. 2.1b. Clearly, the undeformed surfaces overlap whenever

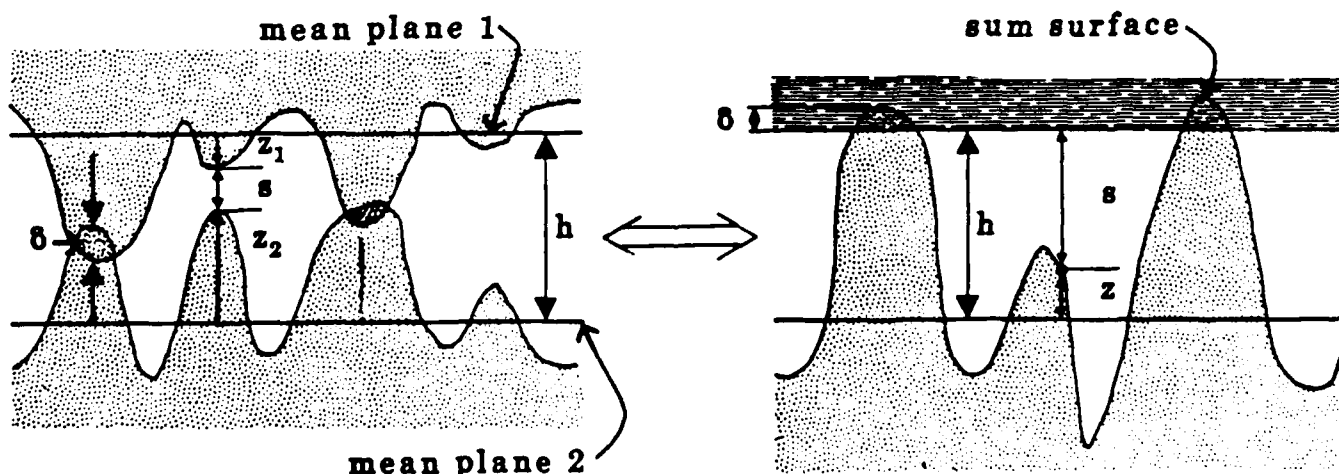
$$s(x,y) < 0$$

As the normal load pressing the surfaces together increases, the approach a decreases and at each minimum of the function s a microcontact nucleates and expands due to local deformation of the surfaces.

It is well known in tribology that techniques used to produce engineering surfaces usually produce a Gaussian distribution of the surface heights z_i . Moreover, the sum z of two Gaussian surfaces is also Gaussian; indeed, TALLIAN [1972] points out that if z_1 and z_2 are not exactly Gaussian, their sum surface will be closer to Gaussian than either surface. If the shape of an asperity is assumed to be paraboloidal, as have been done by several authors, then the peak heights and curvatures are correlated random variables, with the result that a Gaussian distribution of heights and curvatures may lead to a cumulative probability distribution of surface heights which is non-Gaussian. As noted earlier, this issue has been studied by HISAKADO [1974], HISAKADO and TSUKIZOE [1974], and by FRANCIS [1977], who asserts that if the peak shape is paraboloidal only at its vertex, then the ensemble of peaks can be made to conform to the Gaussian distribution.



(a)



(b)

Figure 2.2 a) Profiles of opposing rough surfaces and b) microtopography of surfaces and the equivalent sum surface z .

To fix ideas, we note that for a given asperity profile, one defines for the random variable $z = z(x,y)$ (the surface height), the autocorrelation function

$$C(x,y) = \lim_{\substack{a \rightarrow \infty \\ b \rightarrow \infty}} \frac{1}{ab} \int_0^a \int_0^b z(X,Y) z(X+x, Y+y) dX dY$$

and the power spectral density $P(u,v)$ as its Fourier transform,

$$P(u,v) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} C(x,y) \exp[-i(xu + yv)] dx dy$$

The power spectral moments are

$$m_{ij} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(u,v) u^i v^j du dv$$

and the r.m.s. roughness σ is the variance,

$$\sigma^2 = C(0,0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(u,v) du dv$$

If the standard deviations of z , $\partial z / \partial x$, $\partial^2 z / \partial x^2$ are σ , σ' , σ'' , respectively, then the spectral width parameter α is defined by $\alpha^{1/2} = \sigma \sigma'' / \sigma'^2$.

With these notations in hand, we can define the joint probability density function for random surface peak heights and curvatures is (Cf NAYAK [1971]).

$$f_{\xi\eta}(\xi, \eta) = \frac{1}{\pi} \left(\frac{3}{1-\beta^2} \right)^{1/2} [\xi^2 - 1 + \exp(-\xi^2)] \times \exp \left[-\frac{1}{2(1-\beta^2)} (\eta - \xi)^2 \right] \quad (2.2)$$

where $\beta = (3/2\alpha)^{1/2}$, and ξ and η are non-dimensionalized peak heights and curvatures

($\xi = z_{\text{peak}} / \sigma$; $\eta = \kappa\sqrt{3} / \sigma''\sqrt{2}$).

Let

N_p = the number of sum peaks within a nominal contact area A_0

v = area density of asperity peaks.

Then the number of peaks in A_0 is

$$N_p = v A_0$$

Let τ denote the non-dimensional separation,

$$\tau = h / \sigma$$

Then, if X is a generic variable of interest associated with the micromechanics of an asperity, X being a function of ξ and η , then its *expected value* is

$$E(X) = \int_{-\infty}^{\infty} \int_0^{\infty} X(\xi - \tau, \eta) f_{\xi\eta}(\xi, \eta) d\eta d\xi \quad (2.3)$$

and the *macrocontact expectation* of X is obtained by summing X over all microcontacts:

$$\chi(\tau) = N_p E(X) = v A_0 \int_{-\infty}^{\infty} \int_0^{\infty} X(\xi - \tau, \eta) f_{\xi\eta}(\xi, \eta) d\eta d\xi \quad (2.4)$$

The notation η' in (2.3) and (2.4) is intended to indicate that if the approach of FRANCIS [1977] is to be used, in which the shape of an asperity is paraboloidal only at its vertex, then, instead of writing X is a function of peak heights and radii R of curvature, we should use the modified radius R' , with, e.g.

$$\eta' = [1 + \omega(\xi - \tau)] \eta$$

ω being a non-analytic function of $\xi - \tau$.

2.2 Numerical Calculation of Macrocontact Expectations. The macrocontact expectation χ in (2.4) is to be evaluated by numerical quadrature. Toward this end, we partition the ξ, η - plane into a mesh Δ of rectangular elements of area $\Delta\xi\Delta\eta$, and approximate χ by

$$\chi_{\Delta}(\tau) = \nu A_0 \sum_{ij=1}^N \bar{X}(\xi_i - \tau, \eta_{ij}) f_{\xi\rho}^{ij} \Delta\xi\Delta\eta \quad (2.5)$$

where

$$\bar{X}(\xi_i - \tau, \eta_{ij}) = \begin{cases} X[\xi_i - \tau, (1 + \omega(\xi_i - \tau)\eta_j)] & \text{if } \xi_i - \tau > 0 \\ 0 & \text{if otherwise} \end{cases} \quad (2.6)$$

and

$$f_{\xi\eta}^{ij} = f_{\xi\eta}(\xi_i, \eta_j) \quad (2.7)$$

Denoting

$$\begin{aligned} X^{ij} &= \bar{X}(\xi_i - \tau, \eta_{ij}) \\ F_{ij} &= f_{\xi\eta}^{ij} \Delta\xi\Delta\eta \end{aligned} \quad (2.8)$$

the approximate macrocontact expectation of X can be written

$$\chi_{\Delta}(\tau) = \nu A_0 \sum_{i,j} X^{ij} F_{ij} . \quad (2.9)$$

Our plan is to return to (2.9) once we have computed the response (particularly, the true microcontact area and pressure) of a typical asperity to an approach history and to use (2.9) to compute an approximate global expected value of the response, representing a macro-mechanical model of the interface.

3. ADHESION

An important contributing factor to friction on contact surfaces is adhesion: the intermolecular attractive forces that depend on atomic spacing and the corresponding surface energies of materials. For highly polished uncontaminated surfaces, adhesion forces can be very large, leading to the virtual welding of one surface to another, while for rough contaminated surfaces, adhesion effects are often negligible. For rough engineering surfaces under common working conditions, adhesion effects can be significant, so that a rational model of contact and friction should take them into account.

The DMT adhesion model, proposed by DERJAGUIN, MULLER and TOPOROV [1975] and later refined by MULLER, DERJAGUIN, and TOPOROV [1983], attempts to characterize the attractive forces on a spherical elastic asperity in contact with a rigid flat, assuming that the shape of the deformed asperity is given by the Hertz theory and that no attractive forces exist in the contact region. The JKR model, due to JOHNSON, KENDALL, and ROBERTS [1971] also analyzes the elastic spherical asperity-flat problem with Hertz theory, but assumes that attractive forces are confined to the contact area and that the attractive forces produce an elastic deformation of the asperity. The JKR model has been found to be more suitable for soft materials, such as rubbers, while the DMT model is claimed to be more suitable for harder materials with high surface energies (see CHANG, ETSION, and BOGY [1986b] and PASHLEY [1984]). Survey papers on developments in adhesion models have been contributed by PASHLEY, PETHICA, and TABOR [1984] and by PASHLEY and PETHICA [1985]. See also MACFARLANE and TABOR [1950a,b].

We shall include adhesion effects in our contact and friction theory by using an approach similar to, but more general than that of the DMT model. We continue to assume that the surfaces are isotropic, rough, and have a Gaussian distribution of peak heights, that there is no interaction between asperities, and that it suffices to consider a single, paraboloidal asperity impinging on a rigid flat. Following MULLER et al [1983], we characterize the *attractive adhesion pressure* $q(s)$ as that attractive force per unit surface area, acting normal to the mean asperity height plane,

resulting from the Lennard-Jones interaction (interatomic) potential Φ of surface physical chemistry. Then,

$$q(s) = \frac{8\Delta\gamma}{3\epsilon} \left[\left(\frac{\epsilon}{s}\right)^3 - \left(\frac{\epsilon}{s}\right)^9 \right] \quad (3.1)$$

where

s = the separation of the two surfaces *outside* the contact area

ϵ = the intermolecular distance (generally $\epsilon = 0.3 - 0.5$ nm).

In (3.1), $\Delta\gamma$ is the surface energy of adhesion and is defined as follows: if γ_1 and γ_2 are the surface energies of surfaces 1 and 2, respectively, before contact, and γ_{12} is the joint surface energy of the interface after contact, then

$$\Delta\gamma = \gamma_1 + \gamma_2 - \gamma_{12} \quad (3.2)$$

For values of surface energies of adhesion for various metals, see RABINOWICZ [1971] or FERRANTE, SMITH, and ROSE [1983].

The situation of interest here is illustrated in Fig. 3.1. The deformation of a typical asperity is brought about by contact, owing to a change in the approach h of the surfaces, and by the exertion of adhesion tractions g on the material surface, which are functions of the distance s between the flat and the deformed asperity. Thus, if v is the displacement in the z -direction of a material particle on the surface,

$$s = h - z - v$$

To determine v , specific material properties of the asperity must be prescribed. Frequently, v is assumed to be small in comparison with z and is neglected. The principal mathematical difficulty inherent in characterizing the adhesion pressure is that it is developed only on surface material outside the contact area, which, a-priori, is unknown. We shall take up the problem of resolving this non-linear boundary-value problem in the next section. Several concluding remarks on adhesion are in order at this point.

Remarks. 3.1 FULLER and TABOR [1975], using the JKR model of adhesion and the rough surface asperity-based model of GREENWOOD and WILLIAMSON [1966], presented an *adhesion parameter* θ of the form

$$\theta = \frac{E\sigma}{\Delta\gamma} \sqrt{\frac{\sigma}{R}} \quad (3.3)$$

where E is the effective modulus of elasticity of the contact surfaces

$$E^{-1} = (1 - \nu_1^2) E_1^{-1} + (1 - \nu_2^2) E_2^{-1}$$

E_i, ν_i being the Young's modulus and Poisson's ratio of surface i , σ is the standard deviation of surface heights, and R the mean radius of spherical elastic asperities. The larger the value of θ , the less significant the effects of adhesion, and for rubber spheres, experiments showed that adhesion become negligible for $\theta > 10$.

3.2 CHANG, ETSION, and BOGY [1986b] presented a study of adhesion effects using the DMT model and an elasto-plastic asperity model of the contact surface. They investigated the importance of adhesion with varying values of surface energy and plasticity, as measured by the *plasticity index* of GREENWOOD and WILLIAMSON [1966],

$$\psi = \frac{2E}{\pi KH} \sqrt{\frac{\sigma}{R}} \quad (3.4)$$

where H is the Brinell hardness (of the softer of the two materials) and $K = 0.454 + 0.41\nu_1$, ν_1 being the Poisson ratio of this material. These authors concluded that the "pull-off force" due to adhesion (i.e., the integrated suction force due to adhesion) becomes negligible for hard steel when the adhesion index θ is greater than 100, as compared to $\theta > 10$ for rubber; the adhesion force is

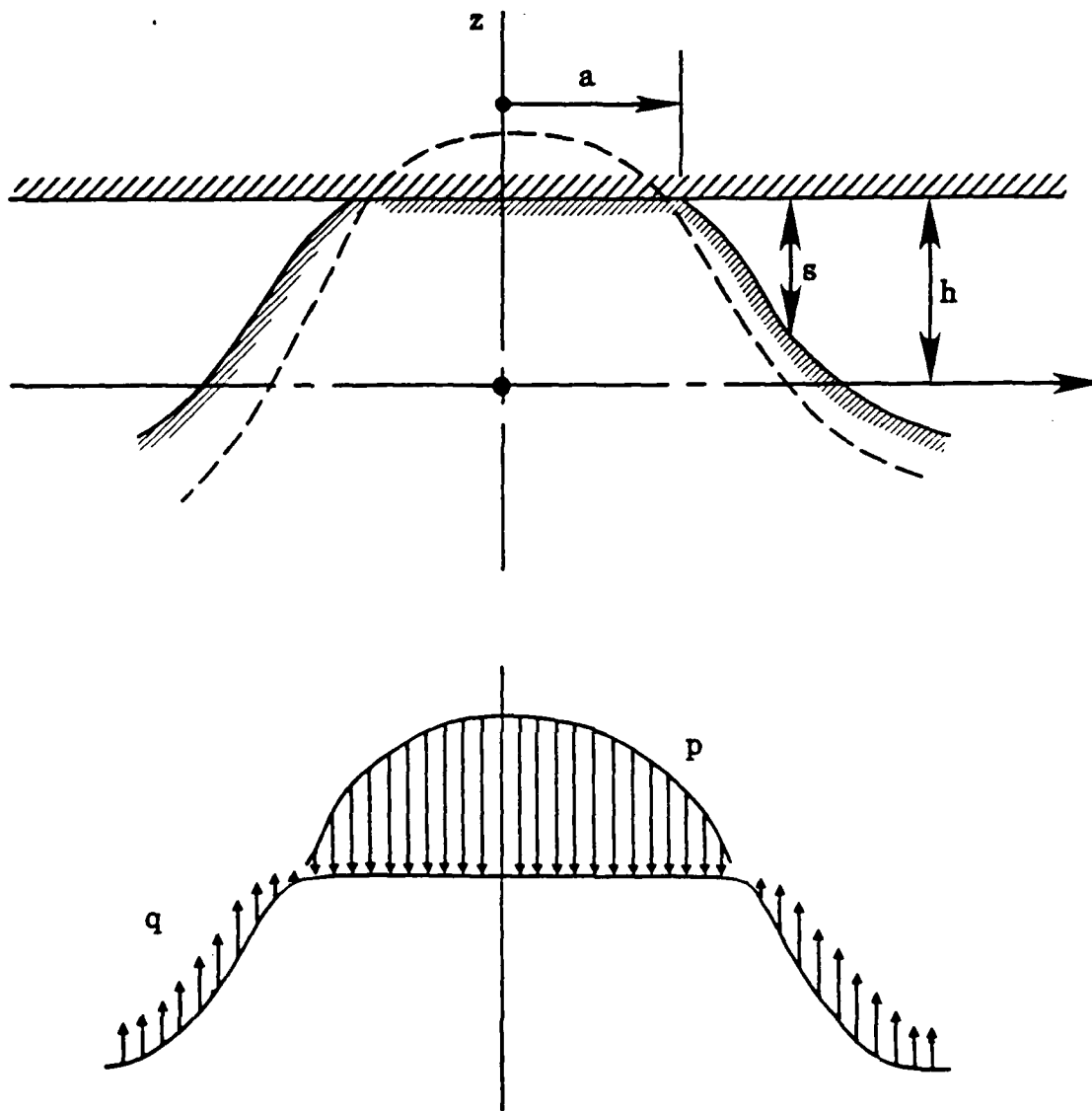


Figure 3.1 Pressures on a deformed asperity; the adhesion pressures q due to molecular attraction and the pressure p on the contact surface.

negligible compared to the contact load when the plasticity index $\psi \geq 2.5$ or when the surface energy $\Delta\gamma \leq 0.5 \text{ J/m}^2$ for a sufficiently small external force. They concluded that "for smooth clean surfaces the adhesion can be well over 20 percent of the contact load and, thus, cannot be neglected."

3.3 Adhesion forces are time dependent and generally increase with time of contact, eventually acquiring a constant value for static contacts. Thus, the static adhesion models generally attempt to predict the maximum value of adhesion forces and can overestimate adhesion effects for dynamic contacts.

4. CONTACT MECHANICS OF A VISCOPLASTIC ASPERITY

We now focus on the analysis of a typical asperity in contact with rigid flat. The asperity is a body of revolution, symmetric about its $z = x_3$ - axis, and subjected to adhesion pressures q on its exterior surfaces that are not in contact with a rigid flat, and to contact pressures due to its indentation into the rigid flat. We shall assume that the asperity is composed of an isotropic viscoplastic material characterized by a collection of constitutive equations of the internal state-variable type. The governing equations in cartesian coordinates are listed as follows:

1. *Momentum Equations.* Considering quasi-static deformations for now, we have

$$\sigma_{ij,j} = 0 \quad (4.1)$$

where

σ_{ij} = the Cauchy stress tensor at a point $x = (x_1, x_2, x_3) \in \Omega$, Ω being the open material domain of the asperity

and $\sigma_{ij,j}$ is the divergence of the stress rate $\dot{\sigma}_{ij}$.

2. *Strain-Rates.* The rates $\dot{\epsilon}_{ij}$ of the total infinitesimal strains are given in terms of the

velocity gradients u_{ij} by the usual relation

$$\begin{aligned}\epsilon_{ij} &= \frac{1}{2} (u_{i,j} + u_{j,i}) \\ &= \epsilon_{ij}^e + \epsilon_{ij}^n\end{aligned}\quad (4.2)$$

where ϵ_{ij}^e and ϵ_{ij}^n are the elastic and inelastic strain rates, respectively.

3. *Constitutive Relations.* A quite broad class of materials is characterized by constitutive equations of the form

$$\begin{aligned}\sigma_{ij} &= E_{ijkl} \epsilon_{kl}^e \\ \epsilon_{ij}^n &= f_{ij}(\sigma, z) \\ z_k &= g_k(\sigma, z)\end{aligned}\quad (4.3)$$

where E_{ijkl} is Hooke's tensor of elasticities, f_{ij} and g_k are smooth functions of the stress σ and of z , and $z = (z_k)$ is a vector- or scalar-valued internal state variable characterizing various micromechanical changes in the material. Constitutive models of this type can be used to characterize viscoelastic, elasti-plastic, and viscoplastic response of metals by appropriate choices of the constitutive functionals; eight theories of this form, proposed by theoretical and experimental studies of different authors on viscoplasticity of metals, are summarized in BASS [1985]. As one example, we mention the theory of BODNER and PARTOM [1972] and BODNER and STOUFFER [1983], for which

$$\begin{aligned}\epsilon_{ij}^n &= f_{ij}(\sigma) \\ &= \frac{D_0 \sigma_{ij}}{J^{1/2}} \exp \left[-\frac{1}{2} (H_{kl} H_{lk} / 3J_2)^m \frac{m+1}{m} \right] \\ z &= f_{ij}(\sigma) \sigma_{ij}'\end{aligned}\quad (4.4)$$

Here $\sigma_{ij} = \sigma_{ij} - \delta_{ij} \sigma_{kk} / 3$ is the deviatoric stress, J_2 is the second principal invariant of σ_{ij} , D_0 and m are material parameters and H_{kl} is a hardness tensor. In this case, z is scalar-valued and coincides with the "plastic work": $z = \epsilon_{ij}^n \sigma_{ij}$. The parameter m is related to the strain-rate sensitivity of the material. If the material response is such that the nonelastic strains ϵ_{ij}^n are not associated with volume changes, ϵ_{ij}^n can be replaced by its deviatoric part, $\epsilon_{ij}^{n'}$. This type of theory can be used to model strain hardening, cyclic stress-strain relations for rate-dependent plasticity, and other inelastic phenomena. Figure 4.2 contains a cyclic stress-strain curve for titanium calculated using the constitutive equations (4.4).

4. *Boundary and Unilateral Contact Conditions*. The asperity can be viewed as a paraboloidal proturbance of a deformable half-space. The condition that it cannot penetrate the rigid flat is

$$u_3 - s \leq 0 \quad \text{on } \Gamma \quad (4.5)$$

where $s = h - z$ is the distance from the asperity surface Γ to the flat. The contact region Γ_c is defined by

$$\Gamma_c = \{ \mathbf{x} \in \Gamma \mid u_3(\mathbf{x}) = s(\mathbf{x}) \}$$

Outside Γ_c we have the traction boundary condition

$$\sigma_{3j} n_j = q; \quad \sigma_{1j} n_j = \sigma_{2j} n_j = 0 \quad \text{on } \Gamma - \Gamma_c \quad (4.6)$$

where $q = q(s)$ is the adhesion pressure defined by (3.1). Elsewhere, we have

$$u_i(\mathbf{x}) \rightarrow 0 \quad \text{as } |\mathbf{x}| \rightarrow +\infty$$

5. *Initial Conditions*. Smooth functions $\mathbf{u}^0(\mathbf{x})$ and $z^0(\mathbf{x})$ are prescribed such that for $\mathbf{x} \in \Omega$,

$$u_i(\mathbf{x}, 0) = u_i^0(\mathbf{x}); \quad z(\mathbf{x}, 0) = z^0(\mathbf{x}) \quad (4.7)$$

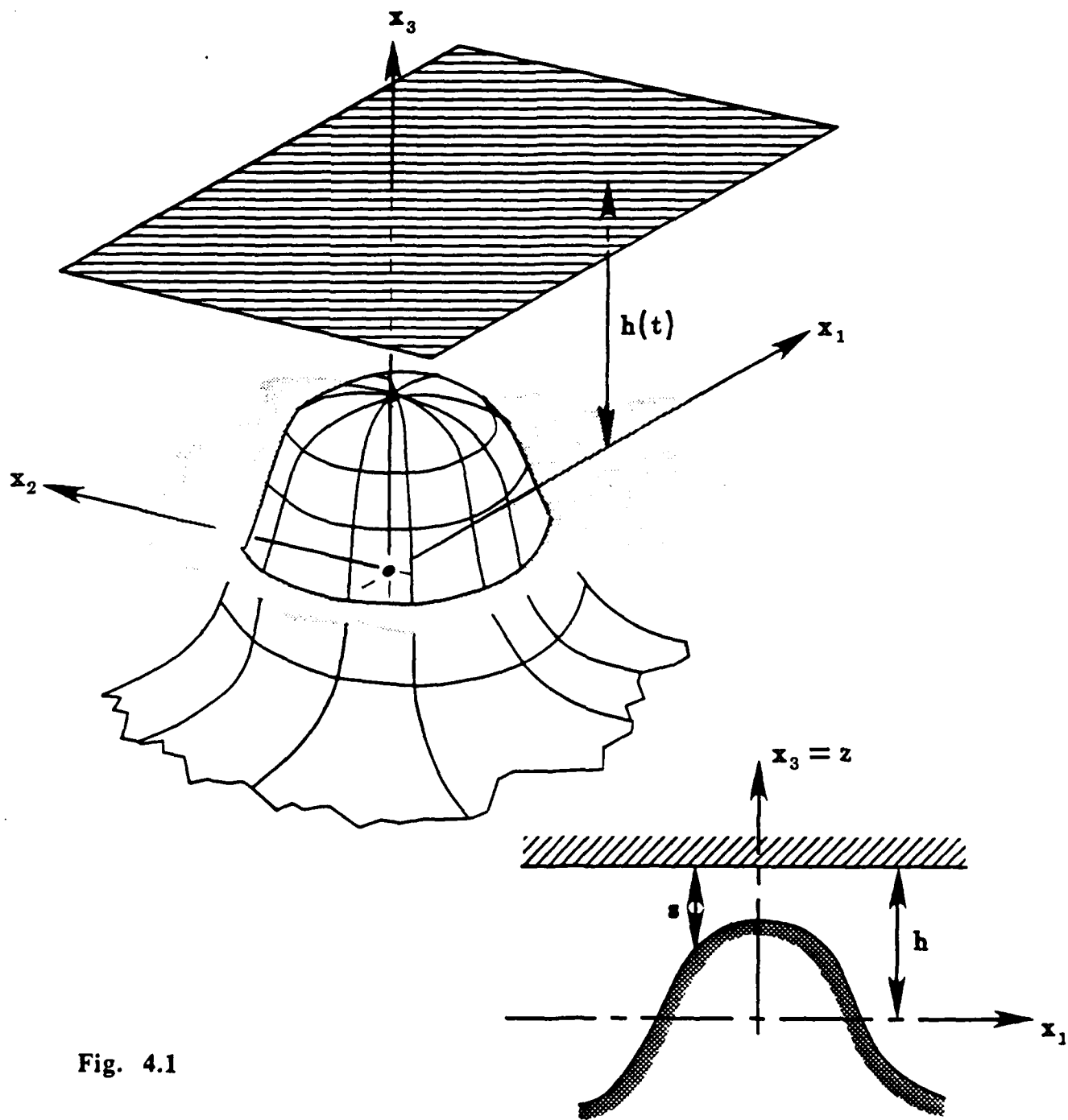


Fig. 4.1

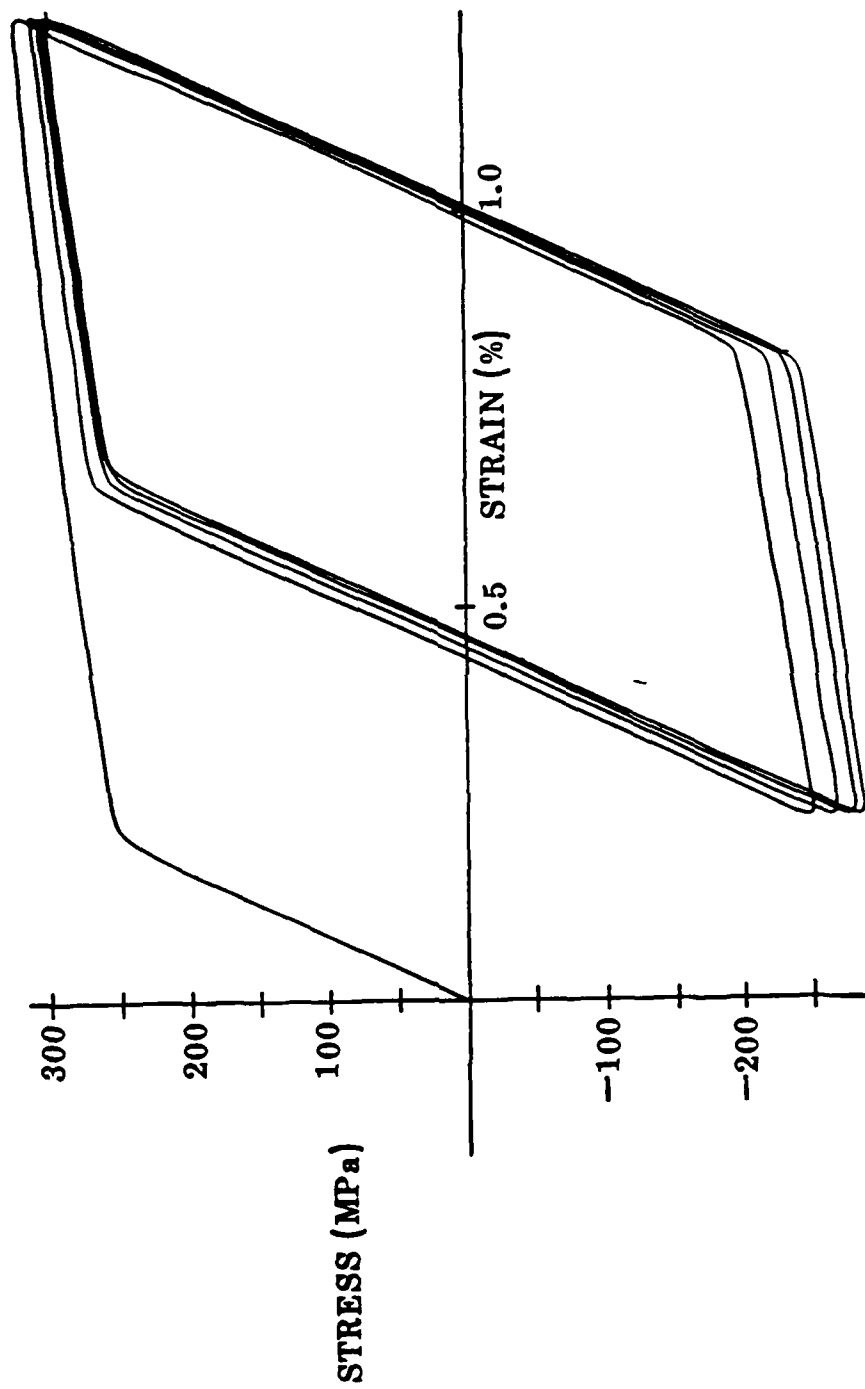


Fig. 4.2 Computed Stress-Strain Curves for Titanium for Strain Controlled Cycling with Positive Strain Limits showing Creep Relaxation (Bodner's Equations).

4.1 Variational Formulation. It is useful to construct a weak or variational formulation of the initial-boundary-value problem described in the preceding subsection. We begin by introducing the space V of admissible displacements and Z of admissible variations in the internal state variable z ; e.g.

$$V = \{ v = (v_1, v_2, v_3) \mid v_i \in W^{1,p}(\Omega), 1 \leq p \leq \infty, v_i(x) \rightarrow 0 \text{ as } \|x\|_{\mathbb{R}^3} \rightarrow \infty \}$$

$$Z = \{ w = (w_1, w_2, w_3) \mid w_i \in L^q(\Omega), 1 \leq q \leq \omega \} \quad (4.8)$$

where $W^{1,p}(\Omega)$ is the Sobolev space of order $1, p$. Of course, the specific definition of V and Z is problem dependent, in the sense that for each time $t \in [0, T]$ the displacements u_i and state variables z_i will take on values in spaces determined by the particular constitutive functionals used in (4.3) to characterize the material. We also introduce the unilateral constraint set K and the pivot space H :

$$K = \{ v \in V \mid v_3 - s \leq 0 \text{ a.e. on } \Gamma \}$$

$$H = (L^2(\Omega))^3$$

where Γ is the boundary surface of the asperity.

If u_{ij} , σ_{ij} , and $z_i \in L^\infty(0, T, H)$, we have a.e.,

$$u_i(x, t) = u_i^0(x) + \int_0^t u_i(x, t) \, dt$$

$$\sigma_{ij}(x, t) = \sigma_{ij}(x, 0) + \int_0^t \sigma_{ij}(x, t) \, dt \quad (4.9)$$

$$z_i(x, t) = z_i^0(x) + \int_0^t z_i(x, t) \, dt$$

If $u \in L^\infty(0, T, V)$ is the actual displacement field existing in the body and v is an arbitrary test function in the constraint set K , then a simple application of Green's formula yields

$$\int_0^t \int_\Omega \sigma_{ij} (v_{ij} - u_{ij}) \, dx \, dt$$

$$= \int_{\Omega} -\sigma_{ij}(\bullet, 0) (v_{ij} - u_{ij}) dx + \int_{\Gamma} \sigma_{ij} n_j (v_i - u_i) ds \quad (4.10)$$

since $\sigma_{ij,j} = 0$. The *initial stress* is given by σ_{ij} .

$\sigma_{ij}(x, 0) = E_{ijkl} u_{k,l}$. Likewise, from (4.7)

$$\begin{aligned} \int_{\Gamma} \sigma_{ij} n_j (v_i - u_i) ds &= \\ &= \int_{\Gamma-\Gamma_c} q(\bullet, h) (v_3 - u_3) ds + \int_{\Gamma_c} \sigma_{3j} n_j (v_3 - u_3) ds \end{aligned}$$

when q is given by (3.1) with $s = h - z$.

Or, since on Γ_c , $u_3 - s = 0$, we have

$$\begin{aligned} \sigma_{3j} n_j (v_3 - u_3) &= \sigma_{3j} n_j (v_3 - s) + \sigma_{3j} n_j (u_3 - s) \\ &= \sigma_{3j} n_j (v_3 - s) \\ &\geq 0 \quad \forall v \in K \end{aligned}$$

Collecting these results and using (4.9) and (4.10), we are finally led to the following nonlinear initial-boundary-value problem characterized by a *system of variational inequalities of evolution*:

Given initial data u^0, z^0 in (4.7), find $(u, z) \in L^1(0, T, K) \times L^1(0, T, W)$

such that for a.e. $t \in [0, T]$,

$$\begin{aligned} &\int_0^T \int_{\Omega} E_{ijkl} [u_{k,l} - f_{kl}(\sigma(\bullet, t), z(\bullet, t))] (v_{ij} - u_{ij}) dx dt \\ &\geq \int_{\Omega} -\sigma_{ij}^0 (v_{ij} - u_{ij}(t)) dx \\ &\quad + \int_{\Gamma-\Gamma(h)} q(\bullet, h(t)) (v_3 - u_3(t)) ds \end{aligned}$$

c

$$\forall v \in K \quad (4.11)$$

$$\begin{aligned} \int_{\Omega} z_i(t) w_i \, dx &= \int_{\Omega} z_i^0 w_i \, dx \\ &+ \int_0^T \int_{\Omega} g_i(\sigma(\bullet, t), z(\bullet, t)) \, dx \, dt \end{aligned} \quad \forall w \in W \quad (4.12)$$

Here, and throughout this purely formal development, it is understood that the regularity of u and z (particularly the spaces L^r and L^s in (4.11), (4.12)) are determined by the regularity of the data and the domain and by properties of the constitutive equations (4.3). Also the separation h , at this point, is regarded as a parameter dependent on time; i.e., we wish to solve (4.11) and (4.12) for a given *time history* of the separation.

4.2 Numerical Analysis. Owing to the highly nonlinear character of the variational inequalities (4.11), (4.12), it is necessary to employ numerical techniques to solve this problem. Toward this end, we have developed finite element models of these nonlinear evolution equations and a stable time marching scheme to integrate in time the resulting system of nonlinear differential inequalities for the case in which the material is described by the BODNER-PARTOM constitutive equations of (4.4). These schemes allow us to determine the net microcontact area $A(h, t)$ and the microcontact normal force $P(h, t)$ as a function of h for given separation histories and material parameters. Details of these calculations are to be given in a later report.

5. SLIDING RESISTANCE

We now come to a point in the analysis at which it is necessary to characterize the resistance of the rough interface to sliding (i.e., to tangential motions of the reference planes relative to one another.) While this aspect of the modeling approach still requires much study, there appears to be at least three methods available for this purpose. First, BOWDEN and TABOR [1950] estimate the resistance to impending motion (more precisely, the static coefficient of friction) by calculating the shear strength of metallic oxides junction developed on the contact surface. Similarly, CHANG, ETSION, and BOGY [1986c] calculate the tangential load required to reach the fracture strength of metallic junctions as an indication of that tangential force required to produce sliding. VILLIAGGIO [1979], on the other hand, studied the problem of contact of periodically spaced elastic asperities and defined the load at which sliding initiates as that which reduces to zero the curvature of the resisting elastic asperities. FRANCIS [1977] modeled the micro sliding resistance using empirical relations based on existing experimental data.

Extensive sequences of experiments on sliding resistance of thin films, involving 27 different materials and a wide range of normal loads, are described in the papers of BOYD and ROBERTSON [1945], BRISCOE, SCRUTON, and WILLIS [1973], TOWLE [1974], and BRISCOE and TABOR [1975]. In all of these studies, it was discovered that (on a microcontact interface) the interfacial shear stress τ_s during sliding was a function of the normal stress

$\sigma_n = \sigma_{ij} n_j n_i$, and, according to FRANCIS [1977], their empirical findings suggest that

$$\tau_s = \begin{cases} c_1 + c_2 \sigma_n & \text{for light loads} \\ c_3 \sigma_n^m & \text{for intermediate loads} \\ \log^{-1} [c_5 + c(\sigma_n) \log \sigma_n] & \text{for heavy loads} \end{cases} \quad (0.6 < m < 1.4) \quad (5.1)$$

$$0.0 \leq c(\sigma_n) < 1.9; \quad \frac{d(\log \tau_s)}{d(\log \sigma_n)} \geq 0 \quad \text{for heavy loads}$$

wherein loads were varied over a factor of 10 or more and starting from 15 MPa (light), 40 MPa (intermediate) and 200 MPa (heavy). FRANCIS [1977] points out that a good approximation to all of these cases is the simple quadratic function,

$$\tau_s = c_0 + c_1 \sigma_n + c_2 \sigma_n^2 \quad (5.2)$$

where c_0, c_1, c_2 are material constants.

Once a micro-shear resistance is characterized, the macro sliding resistance can be computed using the statistical summation procedures described earlier. To cite one interesting case, if the asperity is a purely elastic sphere of radius R and modulus E , if (5.2) holds, and if the true microcontact area is given by the Hertz solution, e.g.

$$A\left(\frac{h}{\sigma}\right) = \beta^{-1} A_0 \Delta \xi \Delta \eta \sum_{i,j} f_{ij}(\xi_i, \eta_j) / 4\sqrt{3} \quad (5.3)$$

then the *time-averaged expected value* of the coefficient of friction μ can be calculated; viz.

$$\mu\left(\frac{h}{\sigma}\right) = \frac{\sum_{i,j} F_{ij}^N N_{ij}}{W\left(\frac{h}{\sigma}\right)} \quad (5.4)$$

wherein

$$\begin{aligned} W\left(\frac{h}{\sigma}\right) &= \text{macrocontact normal force} \\ &= \frac{\sigma' E A_0}{6\pi\sqrt{3}} \cdot \beta^{3/2} \Delta \xi \Delta \eta \sum_{i,j} \xi_i^{3/2} \eta_j^{-1/2} f_{ij}(\xi_i, \eta_j) \end{aligned}$$

$$F_{ij}^N = A_{ij} [c_0 + c_1 \sigma_n \left(\frac{h}{\sigma}\right) + c_2 \sigma_n \left(\frac{h}{\sigma}\right)^2]$$

Here A_{ij} is the microcontact area at grid point (i,j) and σ_n is the microcontact normal stress given by the classical Hertz theory.

The establishment of the micromechanical normal response in Section 4 due to adhesion and to changes in the approach and the establishment of the micromechanical sliding resistance by (5.1) or (5.2) completes the description of the principal components of the friction model. Specific cases must now be defined and analyzed before the study can continue.

6. FUTURE WORK

The discussion in Sections 1 - 6 of Chapter 2 of this report clearly outlines the general approach that has been taken for friction modeling. In the immediate future, the following issues are to be the focus of our work:

1) *Completion of the Contact Code and Time-Marching Algorithm.* Work is underway to develop an axisymmetric finite element scheme for solving the nonlinear unilateral contact problem (4.11), (4.12) for cyclic loading and damage of a viscoplastic asperity. This is a difficult and time-consuming step in the analysis.

2) *Development of Micromodels.* Once the microanalysis is complete, the results are to be summed in the manner described herein to yield the micromechanical models of the interface. Such calculations should yield a wealth of results, including estimates of the evolution of the coefficient of friction, of damage, etc.

3) *Generalizations in Structural Dynamics.* Once new constitutive models for interfaces are in hand, their use in modeling frictional resistance in structural dynamics can be investigated. Issues relevant to such studies are the existence, regularity, and stability of solutions, the optimal control of the dynamic response of structures with such frictional constraints, the effects of changes in the friction characteristics due to plastic deformation, damage and wear. Finally, applications to representative problems in structural dynamics shall be considered.

CHAPTER 5

ON THE CONTROLLABILITY OF DISTRIBUTED VISCOELASTIC SYSTEMS WITH FRICTIONAL FORCES

1. INTRODUCTION

The dynamics and control of distributed systems subjected to frictional resistive forces has been an open problem in partial differential equations for many years, largely owing to the absence of a meaningful existence theory for dynamical systems with friction. The mathematical difficulties associated with Coulomb friction in problems of elastodynamics were pointed out in 1972 by Duvaut and Lions [1]) and have been the subject of much study in the intervening years. The finite-dimensional case has proved somewhat more tractable, and conditions for the existence of solutions of discrete dynamical problems with friction have been recently reported by Lötstedt [2] and Jean and Pratt [3]. The complexity of the problem can be appreciated by reviewing the work on dynamics of systems with frictionless contacts by Schatzman [4], Carrero and Pascali [5], Lötstedt [6], Amerio and Prouse [7], Amerio [8], and others.

In recent papers, Oden and Martins [9,10] pointed out that one of the principal sources of mathematical difficulty was the definition of frictional stresses on the contact surfaces characterized by Coulomb's law. However, an overwhelming volume of experimental data accumulated over a half century suggests that this law is inadequate for modeling actual contacts and resistive forces on deformable bodies. By characterizing the actual normal compliance of elastic interfaces, a constitutive equation for an interface can be developed which yields results in agreement with a sizable collection of experimental results on static and dynamic friction [9]. Moreover, the use of such interface constitutive laws in mathematical models of elastostatics, elastodynamics, and viscoelastodynamics problems with friction produces a tractable theory: results on the existence and local uniqueness of solutions to static problems in elasticity with Oden-Martins [9] nonlinear friction laws were recently established by Rabier, et al [11] and of dynamics problems on elasticity and viscoelasticity by Martins and Oden [10]. These new theories and results set the stage for the study of the optimal control of such systems, taken up in the present paper.

In the present study, we establish the existence of a class of optimal controls for a broad class of problems in the dynamics of linear viscoelastic bodies with contact and friction laws of the type introduced by Oden and Martins [9]. We are able to show, under very mild restrictions on the data, that for a control of the form $\beta(t)H$, H being a functional characterizing all external forces on the system, a $\beta \in H^1(0,T)$ can be found which minimizes a functional relating the L^2 -distance between solutions to the viscoelastodynamics problem and an arbitrary target motion $z(\cdot,t)$ over a time interval $[0,T]$.

Following this Introduction, we introduce in Section 2 notations and record preliminary results on the formulation of the contact problem in viscoelastodynamics. In Section 3, a weak or

variational formulation of the problem is presented and an existence result from [10] is recorded. The optimal control problem is introduced in Section 4 and the major results are established in Section 5.

We believe that these results are the first of this type on optimal control of distributed-parameter systems in contact with friction.

2. PRELIMINARIES

We begin by considering a metallic body, the interior of which is a bounded open domain Ω in \mathbb{R}^N ($N = 2$ or 3). The boundary Γ of Ω is smooth (at least Lipschitz continuous) contains open subsets Γ_D , Γ_F , and Γ_C such that $\bar{\Gamma} = \bar{\Gamma}_D \cup \bar{\Gamma}_F \cup \bar{\Gamma}_C$, $\text{meas}(\bar{\Gamma}_\alpha \setminus \bar{\Gamma}_\alpha) = 0$, $\alpha \in \{D, F, C\}$. Material particles (points) in Ω with cartesian coordinates x_i , $1 \leq i \leq N$, are denoted x , the volume measure by dx , and the surface area measure by ds , and a unit exterior vector normal to Γ by n .

We shall assume that the body is composed of a linear viscoelastic material, the mechanical response of which is characterized by the constitutive law,

$$\sigma_{ij} = E_{ijkl} u_{k,l} + C_{ijkl} \dot{u}_{k,l} \quad 1 \leq i, j, k, l \leq N \quad (2.1)$$

where σ_{ij} ($= \sigma_{ji}$) are the cartesian components of the Cauchy stress tensor, $E_{ijkl} = E_{ijkl}(x)$ and $C_{ijkl} = C_{ijkl}(x)$ are the arrays of elastic and viscoelastic parameters at point $x \in \Omega$, $u_k = u_k(x, t)$ are the components of displacement at x at time $t \in [0, T] \subset \mathbb{R}$, $u_{k,l} = \partial u_k / \partial x_l$ and $(\dot{\cdot})$ denotes differentiation with respect to time. The body is subjected to body forces of intensity b per unit volume, to surface tractions t on Γ_F , and is possibly in contact with a moving neighboring body on the contact surface Γ_C which moves relative to Ω at a velocity \dot{U}_T tangent to Γ_C . The partial differential equations, inequalities, and conditions governing the behavior of the body are listed as follows:

Linear Momentum

$$\sigma_{ij} (u, \dot{u})_{,j} + b_i = \rho \ddot{u}_i \quad \text{in } \Omega \times (0, T) \quad (2.2a)$$

where $\sigma_{ij} (u, \dot{u})$ is the expression on the right-hand side of (2.1).

Boundary Conditions

$$u_i = 0 \quad \text{on } \Gamma_D \times (0,T) \quad (2.2b)$$

$$\sigma_{ij}(u, \dot{u}) n_j = t_i \quad \text{on } \Gamma_F \times (0,T) \quad (2.2c)$$

Contact Interface Conditions

$$\sigma_n(u, \dot{u}) = -C_n h(u_n)^{m_h} \quad (2.2d)$$

$$u_n \leq g \Rightarrow \sigma_T(u, \dot{u}) = 0$$

$$\begin{array}{l}
 u_n > g \left\{ \begin{array}{l}
 |\sigma_T(u, \dot{u})| \leq C_T h(u_n)^{m_T} \\
 \text{and} \\
 |\sigma_T(u, \dot{u})| < C_T h(u_n)^{m_T} \\
 \Rightarrow w_T = 0 \\
 \\
 |\sigma_T(u, \dot{u})| = C_T h(u_n)^{m_T} \\
 \Rightarrow \text{there exists } \lambda > 0 \\
 \text{such that} \\
 w_T = -\lambda \sigma_T(u, u) \\
 \text{on } \Gamma_C \times (0,T)
 \end{array} \right. \quad (2.2e)
 \end{array}$$

Initial Conditions

$$u(x,0) = u_0(x)$$

$$\dot{u}(x,0) = u_1(x)$$

$$x \in \Omega \quad (2.2f)$$

In (2.2a), $\rho = \rho(x)$ is the mass density and in (2.2d) and (2.2e),

$$\begin{aligned}
 \sigma_n(u, \dot{u}) &= \text{the normal stress (contact stress) on } \Gamma_C \\
 &= \sigma_{ij}(u, \dot{u}) n_i n_j
 \end{aligned}$$

$$\begin{aligned} h(u_n) &= \text{the approach of the two contact surfaces} \\ &= (u_n - g)_+ \end{aligned}$$

with $u_n = u_i n_i$ the normal component of displacement, g the initial gap between surfaces, and $(\cdot)_+$ the positive part of the indicated argument ($\phi_+ = \max(0, \phi)$).

$$\begin{aligned} \sigma_{Ti}(u, \dot{u}) &= \text{the tangential (frictional) stress component on } \Gamma_C, \\ &= \sigma_{ij}(u, \dot{u}) n_j - n_i \sigma_n(u, \dot{u}) \end{aligned}$$

$$\begin{aligned} \dot{w}_T &= \text{the relative (slip) velocity of the contact surfaces,} \\ &= \dot{u}_T - \dot{U}_T \end{aligned}$$

where $u_{Ti} = u_i - n_i u_n$.

In (2.2d) and (2.2e), C_n, m_n, C_T, m_T are material constants characterizing the mechanical response of the interface. Equation (2. d) is the power-law constitutive equation for the normal compliance of the interface; (2.2e) defines the corresponding tangential response characteristics. Justification for the use of such interface laws is given in ODEN and MARTINS [9]. These laws correspond to a generalized Coulomb law in which the coefficient of friction is dependent on the deformation and is of the form, $\mu = C|\sigma_n(u, \dot{u})|^\alpha$, with $\alpha = (m_T/m_n) - 1$ and $C = C_T/C_n^{m_n/m_T}$.

3. A VARIATIONAL FORMULATION

We now record a weak formulation of Problem (2.2), in which regularity requirements on u and on the data are relaxed. We introduce the spaces of admissible functions,

$$V = \{ v = (v_1, v_2, \dots, v_N) \in (H^1(\Omega))^N \mid v = 0 \text{ a.e. on } \Gamma_D \} \quad (3.1a)$$

$$H = (L^2(\Omega))^N \quad (3.1b)$$

$$V' = \text{topological dual of } V \quad (3.1c)$$

The space V is equipped with the usual norm,

$$\|v\| = \left\{ \int_{\Omega} (v_{i,j} v_{i,j} + v_i v_i) dx \right\}^{1/2}$$

and (\cdot, \cdot) and $|\cdot|$ denote the L^2 -inner product and norm on H , and $\langle \cdot, \cdot \rangle$ denotes duality pairing on $V' \times V$. Throughout, we assume for simplicity that

$$\rho(x) = 1 \text{ and } \text{meas}(\Gamma_D) > 0. \quad (3.2)$$

In (3.1a), v on the boundary is interpreted in the usual sense of traces of $H^1(\Omega)$ - functions.

The material coefficients in (2.1) are assumed to satisfy the following conditions:

$$\begin{aligned} E_{ijkl}, C_{ijkl} &\in L^\infty(\Omega) \\ E_{ijkl} &= E_{jilk} = E_{ijlk} = E_{klij} \\ C_{ijkl} &= C_{jikl} = C_{ijlk} = C_{klij} \\ \exists \alpha_E, \alpha_C &> 0 \text{ such that} \\ E_{ijkl} A_{kl} A_{ij} &\geq \alpha_E A_{ij} A_{ij}; \quad C_{ijkl} A_{kl} A_{ij} \geq \alpha_C A_{ij} A_{ij} \\ &\text{for every symmetric matrix } A_{ij} \in \mathbb{R}^{N \times N} \end{aligned} \quad (3.3)$$

Whenever (3.3) hold, the bilinear forms $a: V \times V \rightarrow \mathbb{R}$ and $c: V \times V \rightarrow \mathbb{R}$ (representing the internal virtual work of the stress σ_{ij}) defined by

$$a(v, w) = \int_{\Omega} E_{ijkl} v_{k,l} w_{i,j} dx \quad (3.4a)$$

$$C(v, w) = \int_{\Omega} C_{ijkl} v_{k,l} w_{i,j} dx \quad (3.4b)$$

for $v, w \in V$, are continuous and V -elliptic, i.e., there exist positive constants $M_E, M_C, \alpha_E, \alpha_C$ such that for any w, v in V ,

$$|a(w, v)| \leq M_E \|w\| \|v\|, \quad |c(w, v)| \leq M_C \|w\| \|v\| \quad (3.5a)$$

$$a(v,v) \geq \alpha_E \|v\|^2, \quad c(v,v) \geq \alpha_C \|v\|^2 \quad (3.5b)$$

We also assume that

$$1 \leq m_n, m_T \begin{cases} < +\infty & \text{if } N = 2 \\ \leq 3 & \text{if } N = 3 \end{cases} \quad (3.6)$$

We assume that

$$b(t) \in H \quad \text{and} \quad t(t) \in (L^{q'}(\Gamma_C))^N \quad (3.7a)$$

$$C_n, C_T \in L^\infty(\Gamma_C), \quad C_n, C_T \geq 0 \quad \text{a.e. on } \Gamma_C \quad (3.7b)$$

$$g \in L^q(\Gamma_C) \quad (3.7c)$$

where

$$q = 1 + \max(m_n, m_T), \quad q' = q / (q-1)$$

$$b(t) \text{ denotes } x \rightarrow b(x, t); \quad t(t) \text{ denotes } s \rightarrow t(s, t)$$

where $s = (s_1, s_2, \dots, s_N)$ is a point on Γ (in particular Γ_F). Then the work done by the external forces is defined by the functional,

$$\begin{aligned} F(t) &\in V' \\ \langle F(t), v \rangle &= \int_{\Omega} b(t) \cdot v \, dx + \int_{\Gamma_F} t(t) \cdot v \, ds \end{aligned} \quad (3.8)$$

Also, to characterize the prescribed slip velocity \dot{U}_T on Γ_C , we introduce the function

$$\begin{aligned} \Phi(t), \dot{\Phi}(t) &\in V \\ \Phi_n(t), \dot{\Phi}_T(t) &= \dot{U}_T(t) \quad \text{a.e. on } \Gamma_C \end{aligned} \quad (3.9)$$

For discussions of such decompositions of functions in V into tangential and normal components, see KIKUCHI and ODEN [12].

Finally, to characterize the work done by normal and frictional forces on Γ_C , we introduce the nonlinear forms,

$$P : V \rightarrow V'$$

$$\langle P(w), v \rangle = \int_{\Gamma_C} C_n h(w_n)^{m_n} v_n ds \quad (3.10a)$$

$$j : V \times V \rightarrow \mathbb{R}$$

$$j(w, v) = \int_{\Gamma_C} C_T h(w_n)^{m_T} |v_T| ds \quad (3.10b)$$

for arbitrary w, v in V .

We are now ready to state the following weak formulation of the problem.

Find the motion $u(t) : [0, T] \rightarrow V$ such that

$$\begin{aligned} & \langle \ddot{u}(t), v - \dot{u}(t) \rangle + a(u(t), v - u(t)) + C(u(t), v - \dot{u}(t)) \\ & + \langle P(u(t)), v - \dot{u}(t) \rangle + j(u(t), v - \dot{\Phi}(t)) - j(u(t), \dot{u}(t) - \dot{\Phi}(t)) \\ & \geq \langle F(t), v - \dot{u}(t) \rangle \quad \forall v \in V \end{aligned} \quad (3.11)$$

If $\partial_2 j(u(t), \dot{u}(t) - \dot{\Phi}(t))$ denotes the partial subdifferential of the friction functional j with respect to the second argument (i.e., the velocity), then (3.10) can be written in the equivalent form,

$$\begin{aligned} & \text{Find } u(t) : [0, T] \rightarrow V \text{ such that} \\ & F(t) \in \partial_2 j(u(t), \dot{u}(t) - \dot{\Phi}(t)) \\ & + P(u(t)) + C \dot{u}(t) + K u(t) + \ddot{u}(t) \end{aligned} \quad (3.12)$$

where $K, C \in \mathcal{L}(V, V')$ are the operators defined by $\langle K w, v \rangle = a(w, v)$, $\langle C w, v \rangle = c(w, v)$.

We are now in a position to state a major result on the existence and uniqueness of solutions to (3.11).

Theorem 3.1 (MARTINS and ODEN [10]). Let $\text{meas } (\Gamma_D) > 0$, (3.2), (3.5), and (3.6) hold and suppose that

$$\begin{aligned} b &\in L^2(0, T; V') \\ t &\in L^2(0, T; (L^q(\Gamma_F))^N) \\ (\text{so that } F(t) &\in L^2(0, T; V')) \\ u_0 &\in V, u_1 \in H \end{aligned} \tag{3.13}$$

Then there exists a unique solution u to problem (3.11) (or, equivalently, (3.12)) such that

$$\begin{aligned} u &\in L^\infty(0, T; V) \\ \dot{u} &\in L^\infty(0, T; H) \cap L^2(0, T; V') \\ \ddot{u} &\in L^\infty(0, T; V'). \end{aligned} \tag{3.14}$$

This result is obtained using a regularization procedure to smooth the frictional boundary condition along with a standard Galerkin technique. It relies heavily on the compactness of the trace operator from V into appropriate boundary spaces. The condition (3.5), which is used in insuring the appropriate compactness properties, is remarkably confirmed by physical experiments on dry rough surfaces; see [9].

Let us introduce the Hilbert space

$$W = \{ w \in L^2(0, T; V) : \dot{w} \in L^2(0, T; H) \}$$

with norm

$$\|w\|_W = \left(\int_0^T (\|w(t)\|_V^2 + \|\dot{w}(t)\|_H^2) dt \right)^{1/2}$$

An examination of the proof of Theorem 3.1 in [10] reveals the following continuity result.

Theorem 3.2. Let $\text{meas } (\Gamma_D) > 0$, (3.2), (3.5), and (3.6) hold and suppose that $\{F_K\}_{K=1}^{\infty}$ is a sequence in $L^2(0,T;V')$ such that

$$F_K \rightarrow F \text{ weakly in } L^2(0,T;V').$$

Then corresponding sequence of solutions $u(F_K)$ of (3.12) have the property that

$$u(F_K) \rightarrow u(F) \text{ weakly in } W$$

and thus

$$u(F_K) \rightarrow u(F) \text{ in } L^2(0,T;H).$$

Theorems 3.1 and 3.2 provide the basis for a study of the optimal control of distributed systems of the type (3.11). We begin by introducing the notation

$$\begin{aligned} z &\in L^2(0,T;H) \text{ a prescribed target notation for } t \in (0,T), \\ F(t) &\in V' \end{aligned} \quad (3.15a)$$

with

$$F(t) = \beta(t) \Theta, \beta \in H^1(0,T) \quad (3.15b)$$

where Θ is a prescribed functional in V' defined by normalized external forces

$$\langle \Theta, v \rangle = \int_{\Omega} \mathbf{b} \cdot \mathbf{v} \, dx + \int_{\Gamma_F} \mathbf{t} \cdot \mathbf{v} \, ds \quad (3.15c)$$

Here $\beta = \beta(t)$ serves as a control parameter and the functional

$$J : H^1(0,T) \rightarrow \mathbb{R}$$

given by

$$J(\beta) = \left(\int_0^T \|u(\cdot, t; \beta) - z(t)\|_H^2 \, dt + \gamma \|\beta\|_{H^1(0,T)}^2 \right) \quad (3.16a)$$

where $u(\cdot, t; \beta)$ is the solution of (3.12) for the forcing function given in (3.16).

The optimal control problem that we study here is given as follows.

Given $z \in L^2(0,T;H)$ find $\beta_0 \in H^1(0,T)$ (3.16b)
such that

$$J(\beta_0) = \infimum \{ J(\beta) : \beta \in H^1(0,T) \}.$$

The following existence result is a consequence of Theorem 3.2.

Theorem 3.3. Let $(\Gamma_D) > 0$, (3.2), (3.5), (3.6) and (3.16) hold. Then there exists a solution to problem (3.16).

Proof. Let $\{\beta_i\}_{i=1}^\infty$ be a minimizing sequence for (3.16). Then there exists a subsequence and $\{\beta_i\}_{i=1}^\infty$ such that

$$\beta_i \rightarrow \beta_0 \text{ weakly in } H^1(0,T).$$

From Theorem 3.2 we see that

$$u(\beta_i) \rightarrow u(\beta_0) \text{ weakly in } L^2(0,T;H).$$

That β_0 is a solution of (3.16) follows from the weak lower semicontinuity of the mapping

$$\beta \rightarrow \|\beta\|_{H^1(0,T)}.$$

We have established the existence of optimal controls for a class of control problems governed by variational inequalities modeling contract problems with frictional effects. Our goal now is to determine an approximation theory for these problems. Thus, we formulate a class of approximating regularized problems in which the frictional condition is regularized similar to the approach for existence [10]. However, in addition we include a Sobolev regularization term. We determine convergence behavior of the optimal controls for the inclusion of the Sobolev term. This enables us to obtain regularity properties for the optimal controls for regularized problems. These properties allow us to determine limiting behavior for optimal control problems over finite dimensional subspaces of $H^1(0,T)$.

4. REGULARIZED PROBLEMS

In this section we consider the formulation of control problems governed by variational problems that are regularizations of (3.11), c.f. [10]. In contrast to the regularization used in the paper by Oden and Martins to establish the existence of a solution to (3.11), we regularize the

frictional contribution with a smoother function and then introduce a Sobolev regularization as well. This is done to permit us to determine optimality conditions for the optimal controls of the regularized problems. From these conditions we may determine regularity properties of the controls for the regularized problem. We also obtain convergence behavior of the optimal controls as the regularization is allowed to vanish. We begin by introducing the functions

$$\Psi_\varepsilon : \mathbb{R}^N \rightarrow \mathbb{R}, \varepsilon > 0$$

that satisfy the following properties

$$\Psi_\varepsilon \in C^2(\mathbb{R}^N, \mathbb{R}) \text{ for every } \varepsilon > 0 \quad (4.1)(i)$$

$$0 \leq \Psi_\varepsilon(v) \leq |v| \text{ for every } \varepsilon > 0 \text{ and } v \in \mathbb{R}^N \quad (4.1)(ii)$$

$$\Psi_\varepsilon(\theta w + (1 - \theta)v) \leq \theta \Psi_\varepsilon(w) + (1 - \theta) \Psi_\varepsilon(v) \quad (4.1)(iii)$$

$$\text{for every } \varepsilon > 0, (w, v) \in \mathbb{R}^N \times \mathbb{R}^N, \text{ and } \theta \in [0, 1]$$

$$\text{There exists positive constant } D_1 \text{ and a positive function } \varepsilon \rightarrow D(\varepsilon) \quad (4.1)(iv)$$

$$\text{such that for every } \varepsilon > 0 \text{ and } (w, v, u) \in \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^N,$$

$$|\Psi'_\varepsilon(w)(v)| \leq D_1 |v|$$

$$|\Psi''_\varepsilon(w)(v, w)| \leq D_1(\varepsilon) |v| |u|$$

$$\text{There exists a positive constant } D_2 > 0 \text{ such that for every } \varepsilon > 0 \text{ and } v \in \mathbb{R}^N$$

$$|\Psi_\varepsilon(v) - |v|| \leq D_2 \varepsilon$$

We define the regularized functional

$$j_\varepsilon : V \times V \rightarrow \mathbb{R}$$

by

$$j_\varepsilon(w, v) = \int_{\Gamma_C} C_T [(w_n - g)_+]^{m_T} \Psi_\varepsilon(v_T) ds$$

We observe that the derivative with respect to the second argument of $j_\varepsilon(\cdot, \cdot)$ is given by

$$\langle J_\varepsilon(w, v), z \rangle = \langle \partial_v j_\varepsilon(w, v), z \rangle$$

so that

$$\langle J_\varepsilon(w, v), z \rangle = \int_{\Gamma_c} C_T [(w_n - g)_+]^{m_T} \Psi'_\varepsilon(v_T)(z_T) ds \quad (4.2)$$

The variational formulation of the regularized problem of interest in this study is given as the following.

Find a function $t \rightarrow u_\varepsilon(t)$ of $[0, T]$ into V such that for any $v \in V$

$$\begin{aligned} \langle \ddot{u}_\varepsilon(t), v \rangle + \varepsilon a(\ddot{u}_\varepsilon(t), v) + c(u_\varepsilon(t), v) + a(\dot{u}_\varepsilon(t), v) + \langle P(u_\varepsilon(t), v) \\ + \langle J_\varepsilon(u_\varepsilon(t), \dot{u}_\varepsilon(t) - \dot{\Phi}(t)), v \rangle = \langle f(t), v \rangle \end{aligned} \quad (4.3)(i)$$

with initial conditions

$$u_\varepsilon(0) = u_0 \in V \quad (4.3)(ii)$$

$$u_\varepsilon(0) = u_1 \in V$$

and

$$f \in L^2(0, T; V') \quad (4.3)(iii)$$

The proof of the existence of a unique solution to (4.3) is essentially the same as that given in [10]. We use the following convergence result the validity of which follows from a study of the proof of convergence of u_ε to u in [10]. It should be noted that for $f \in B$ a bounded subset of $L^2(0, T; V')$ that the sets

$$\{u_\varepsilon(f) : f \in B, \varepsilon > 0\}$$

and

$$\{u_\varepsilon(f) : f \in B, \varepsilon > 0\}$$

are bounded in $L^\infty(0, T; V)$ and $L^\infty(0, T; H) \cap L^2(0, T; V)$, respectively.

Proposition 4.1. Let $\text{meas}(\Gamma_D) > 0$, (3.2), (3.5), (3.6) and (4.1) hold and suppose that $f_i \rightarrow f$ weakly in $L^2(0, T; V')$ and $\varepsilon_i \rightarrow 0$ as $i \rightarrow \infty$. Then

$$u_{\varepsilon_i}(f_i) \rightarrow u(f) \text{ weak } * \text{ in } \tilde{L}^\infty(0, T; V)$$

$$u_{\varepsilon_i}(f_i) \rightarrow u(f) \text{ weak } * \text{ in } \tilde{L}^\infty(0, T; H) \text{ and weakly in } L^2(0, T; V).$$

Along similar lines but with $\varepsilon > 0$ fixed, we have the continuity result similar to Theorem 3.2.

Proposition 4.2. Let $\varepsilon > 0$ be fixed and let $\text{meas}(\Gamma_D) > 0$, (3.2), (3.5), (3.6) and (4.1) hold. If $f_n \rightarrow f$ weakly in $L^2(0, T; V')$, then $u_\varepsilon(f_n) \rightarrow u_\varepsilon(f)$ weakly in W as $n \rightarrow \infty$.

The control problem for these regularized problems we give as follows.

Find $\beta_\varepsilon \in H^1(0, T)$ such that

$$J^\varepsilon(\beta_\varepsilon) = \inf \{ J^\varepsilon(\beta) : \beta \in H^1(0, T) \} \quad (4.4)(a)$$

where

$$J^\varepsilon(\beta) = \int_0^T \|u_\varepsilon(\cdot, t; \beta) - z(t)\|_H^2 dt + \gamma \|\beta\|_{H^1(0, T)}^2. \quad (4.4)(b)$$

As a consequence of Proposition 4.2, we have the result.

Corollary 4.1. Let the assumption of Proposition 4.2 hold. There exists a solution β_ε to problem (4.4).

Remark 4.1. For β fixed we note that

$$J^\varepsilon(\beta) \geq J^\varepsilon(\beta_\varepsilon).$$

Furthermore, since from Proposition 4.2 we see that $u_\varepsilon(\beta) \rightarrow u(\beta)$ in $L^2(0, T; H)$, it follows that

$$J^\varepsilon(\beta) \rightarrow J(\beta)$$

as $\varepsilon \rightarrow 0$. The set

$$B = \{ \beta_\varepsilon : \varepsilon > 0 \}$$

of optimal controls of (4.4) is bounded in $H^1(0, T)$ and thus is weakly precompact in $H^1(0, T)$.

Theorem 4.2. Let $\text{meas}(\Gamma_D) > 0$, (3.2), (3.5), (3.6) and (4.1) hold and let $\varepsilon_i \rightarrow 0$ as $i \rightarrow \infty$. Then any weak $H^1(0, T)$ limit point β_0 of the sequence $\{\beta_{\varepsilon_i}\}_{i=1}^\infty$ is a solution of (3.16).

Proof. By the above remark there exists a subsequence again $\{\beta_{\varepsilon_i}\}_{i=1}^\infty$ such that

$$\beta_{\varepsilon_i} \rightarrow \beta_0 \text{ weakly in } H^1(0, T)$$

as $i \rightarrow \infty$. From Proposition 4.1 it follows that

$$u_{\varepsilon_i}(\beta_{\varepsilon_i}) \rightarrow u(\beta_0) \text{ in } L^2(0, T; H).$$

Thus, we see that

$$\lim J^\varepsilon(\beta_0) \geq \lim J^{\varepsilon_i}(\beta_{\varepsilon_i}) \geq J(\beta_0)$$

as $i \rightarrow \infty$ and β_0 is a solution of (3.16).

5. OPTIMALITY CONDITIONS FOR REGULARIZED PROBLEMS

Having established existence of optimal controls for regularized problems and their convergence properties as $\varepsilon \rightarrow 0$, we now seek to determine their regularity. We begin by stating a simple result that is a consequence of the mean value theorem.

Lemma 5.1. Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be defined by

$$\phi(x) = [(x - g)_+]^m$$

for $m \geq 1$ and $g \in \mathbb{R}$. Then for $x_1, x_2 \in \mathbb{R}$

$$|\phi(x_1) - \phi(x_2)| \leq C(m) (|x_1 - g|^{m-1} + |x_2 - g|^{m-1}) |x_1 - x_2|.$$

Useful estimates for P and j may now be obtained as Corollaries of Lemma 5.1.

Corollary 5.1. Let m_n satisfy (3.6). Then

$$|\langle P(u_1) - P(u_2), v \rangle| \leq C(m_n) \left[\|u_{1N}\|_V^{m_n-1} + \|u_{2N}\|_V^{m_n-1} + \|g\|_{L^{m_n+1}(\Gamma_c)}^{m_n-1} \right] \|u_{1N} - u_{2N}\|_V \|v\|_V$$

Proof. From Lemma 5.1 it follows that

$$\begin{aligned} |\langle P(u_1) - P(u_2), v \rangle| &= \left| \int_{\Gamma_c} C_n \left([(u_{1N} - g)_+]^{m_n} - [(u_{2N} - g)_+]^{m_n} \right) v_N ds \right| \\ &\leq \int_{\Gamma_c} C(m_n) (|u_{1N} - g|^{m_n-1} + |u_{2N} - g|^{m_n-1}) |u_{1N} - u_{2N}| |v_N| ds \\ &\leq C(m_n) \left[\left(\int_{\Gamma_c} |u_{1N} - g|^{m_n+1} ds \right)^{m_n-1/m_n+1} + \left(\int_{\Gamma_c} |u_{2N} - g|^{m_n+1} ds \right)^{m_n-1/m_n+1} \right] \\ &\quad \left(\int_{\Gamma_c} |u_{1N} - g|^{m_n+1} ds \right)^{1/m_n+1} \left(\int_{\Gamma_c} |v_N|^{m_n+1} ds \right)^{1/m_n+1} \end{aligned}$$

Hence, we see that

$$\begin{aligned} |\langle P(u_1) - P(u_2), v \rangle| &\leq C(m_n) \left[\|u_{1N}\|_{L^{m_n+1}(\Gamma_c)}^{m_n-1} + \|u_{2N}\|_{L^{m_n+1}(\Gamma_c)}^{m_n-1} \right. \\ &\quad \left. + \|g\|_{L^{m_n+1}(\Gamma_c)}^{m_n-1} \right] \|u_{1N} - u_{2N}\|_{L^{m_n+1}(\Gamma_c)} \|v_N\|_{L^{m_n+1}(\Gamma_c)} \end{aligned}$$

Recalling that, under the assumption (3.6), $H^{1/2}(\Gamma_c)$ imbeds continuously into $L^q(\Gamma_c)$ and $V=H^1(\Omega)$ maps continuously into $H^{1/2}(\Gamma_c)$, we have the result.

In a similar manner, we have for j ,

Corollary 5.2. Let m_T satisfy (3.6). Then

$$|j(u_1, v) - j(u_2, v)| \leq C(m_T) [\|u_1\|_V^{m_T-1} + \|u_2\|_V^{m_T-1} + \|g\|_{L^{m_n+1}(\Gamma_c)}^{m_T-1}] \|u_1 - u_2\|_V \|v\|_V.$$

Let $h \in L^2(0, T)$ and define

$$w_\delta = (u(\beta + \delta h) - u(\beta)) / \delta \quad \text{for } \delta > 0.$$

We note that $w = w_\delta$ satisfies

$$\begin{aligned} & (\ddot{w}(t), v) + \varepsilon a(\ddot{w}(t), v) + c(\ddot{w}(t), v) + a(w(t), v) + \\ & \quad + 1/\delta \langle P(u(t; \beta + \delta h) - P(u(t; \beta)), v \rangle + \\ & \quad + 1/\delta \{ \langle J(u(t; \beta + \delta h), \dot{u}(t; \beta + \delta h) - \dot{\Phi}(t)), v \rangle - \langle J(\dot{u}(t; \beta), u(t; \beta) - \dot{\Phi}(t)), v \rangle \} \\ & \quad = h(t) \langle \Theta, v \rangle \end{aligned} \tag{5.1}$$

We now use Corollary 5.1 to obtain for $\delta > 0$ the estimate

$$\begin{aligned} & 1/\delta |\langle P(u(t; \beta + \delta h) - P(u(t; \beta)), v \rangle| \leq \\ & \leq C(m_n) (\|u(t; \beta + \delta h)\|_V^{m_n-1} + \|u(t; \beta)\|_V^{m_n-1} + \|g\|_{L^{m_n+1}}^{m_n-1}) \|w(t)\|_V \|v\|_V \end{aligned}$$

and

$$\begin{aligned} & 1/\delta |\langle J(u(t; \beta + \delta h), \dot{u}(t; \beta + \delta h) - \dot{\Phi}(t)), v \rangle - \langle J(u(t; \beta), \dot{u}(t; \beta) - \dot{\Phi}(t)), v \rangle| \\ & = |1/\delta \int_{\Gamma_c} C_T([(u_n)(t; \beta + \delta h) - g]_+)^{m_T} \Psi'(u(t; \beta + \delta h) - \dot{\Phi}(t))(v) - \\ & \quad - [(u_n)(t; \beta) - g]_+^{m_T} \Psi'(\dot{u}(t; \beta) - \dot{\Phi}(t))(v)) ds|. \end{aligned}$$

Now set

$$A = \frac{1}{\delta} \left| \int_{\Gamma_c} C_T [(u_n)(t; \beta + \delta h) - g]_+^{m_T} - [(u(t; \beta) - g)_+]^{m_T} \Psi'(u(t; \beta + \delta h) - \Phi(t))(v) ds \right|$$

and

$$B = \frac{1}{\delta} \left| \int_{\Gamma_c} C_T [(u_n)(t; \beta) - g]_+^{m_T} (\Psi'(\dot{u}(t; \beta + \delta h) - \dot{\Phi}(t)) - \Psi'(u(t; \beta) - \dot{\Phi}(t)))(v) ds \right|$$

From (4.1) (i v) we see that for A we have

$$A \leq D_1 \int_{\Gamma_c} C_T m_T (|u_n(t; \beta + \delta h) - g|^{m_T-1} + |u_n(t; \beta) - g|^{m_T-1}) |w(t)| |v| ds \quad (5.3)$$

and for B

$$B \leq \hat{D}_1(\epsilon) \int_{\Gamma_c} C_T |u_n(t; \beta) - g|^{m_T} |\dot{w}| |v| ds$$

We now make the restriction of assumption (3.6)

$$1 \leq m_T \begin{cases} < +\infty & \text{if } N = 2 \\ \leq 2 & \text{if } N = 3 \end{cases} \quad (5.4)$$

Under the assumption (5.4) we have that

$$\begin{aligned} B &\leq D_1(\epsilon) C_T \left(\int_{\Gamma_c} |u_n(t; \beta) - g|^{m_T+2} ds \right)^{m_T/(m_T+2)} \|\dot{w}\|_{L^{m_n+2}(\Gamma_c)} \|\dot{v}\|_{L^{m_n+2}(\Gamma_c)} \\ &\leq C(\epsilon, m_T) (\|u(\beta)\|_V^{m_T} + \|g\|_{L^{m_n+1}(\Gamma_c)}^{m_T}) \|\dot{w}\|_V \|\dot{v}\|_V \end{aligned} \quad (5.5)$$

Thus, we see that from (5.1) - (5.3) and (5.5),

$$\begin{aligned}
& (\ddot{\mathbf{w}}(t), \mathbf{v}) + \varepsilon a(\dot{\mathbf{w}}(t), \mathbf{v}) + c(\dot{\mathbf{w}}(t), \mathbf{v}) + a(\mathbf{w}(t), \mathbf{v}) \\
& \leq |h(t)| \langle \Theta, \mathbf{v} \rangle + C(\delta) (\|\mathbf{w}(t)\|_V + \|\dot{\mathbf{w}}(t)\|_V) \|\mathbf{v}\|_V
\end{aligned} \tag{5.6}$$

where $C(\delta)$ depends on the parameter of the problem as well as $u(\beta)$, g , and ε and is bounded for $\delta > 0$ in a bounded set. Now setting $\mathbf{v} = \dot{\mathbf{w}}(t)$ we have

$$\begin{aligned}
& d/dt (\|\dot{\mathbf{w}}(t)\|_H^2 + \varepsilon a(\dot{\mathbf{w}}(t), \mathbf{w}(t)) + a(\mathbf{w}(t), \mathbf{w}(t)) + 2C(\dot{\mathbf{w}}(t), \dot{\mathbf{w}}(t))) \\
& \leq 2|h(t)| \langle \Theta, \mathbf{w}(t) \rangle + 2\hat{C}(\delta) (\|\mathbf{w}(t)\|_V + \|\dot{\mathbf{w}}(t)\|_V) \|\dot{\mathbf{w}}(t)\|_V^2 \\
& \leq C_1 |h(t)|^2 + C_2(\delta) (\|\mathbf{w}(t)\|_V^2 + \|\dot{\mathbf{w}}(t)\|_V^2)
\end{aligned} \tag{5.7}$$

We may now obtain from Gronwall's inequality that

$$\|\dot{\mathbf{w}}_{\delta_i}(t)\|_H^2 + \|\dot{\mathbf{w}}_{\delta_i}(t)\|_V^2 + \|\mathbf{w}_{\delta_i}(t)\|_V^2 + \int_0^t \|\dot{\mathbf{w}}_{\delta_i}(\tau)\|_V^2 d\tau \leq C(\delta) \int_0^t |h(\tau)|^2 d\tau \tag{5.7}$$

where $C(\delta)$ is a function of δ sending bounded sets in \mathbb{R}^+ into bounded sets in \mathbb{R}^+ . In a similar manner, using $\mathbf{v} = \mathbf{w}(t)$, we may show boundedness of $\mathbf{w}(t)$ in $L^2(0, T; V)$. Hence, we see that there is a sequence $\delta_i \rightarrow 0$ such that

$$\begin{aligned}
& \mathbf{w}_{\delta_i} \rightarrow \mathbf{w} \text{ weak } * \text{ in } L^\infty(0, T; V) \\
& \dot{\mathbf{w}}_{\delta_i} \rightarrow \mathbf{w} \text{ weak } * \text{ in } L^\infty(0, T; V) \\
& \ddot{\mathbf{w}}_{\delta_i} \rightarrow \mathbf{w} \text{ weakly in } L^2(0, T; V)
\end{aligned} \tag{5.9}$$

With this preparation, we have the following result: We denote by $\partial u_\varepsilon(\beta)(h)$ the variation of $u_\varepsilon(\beta)$ with increment h .

Proposition 5.1. Let $\text{meas}(\Gamma_D) > 0$, (3.2), (3.5), (3.6), (4.1) and (5.4) hold. Then $\mathbf{w} = \partial u_\varepsilon(\beta)(n)$ exists and satisfies the variational equation

$$\begin{aligned}
& (\ddot{w}(t), v) + \varepsilon a(\ddot{w}(t), v) + C(\dot{w}(t), v) + a(w(t), v) + \int_{\Gamma_C} m_n [(u_n - g)_+]^{m_n-1} w_n(t), v_n \, ds + \\
& + \int_{\Gamma_C} \{ C_T [(u_n - g)_+]^{m_T} \Psi''(\dot{u}_T - \dot{\Phi})(w_T, v_T) + \\
& + m_T C_T [(u_n - g)_+]^{m_T-1} w_n \Psi'(\dot{u}_T - \dot{\Phi})(v_T) \} \, ds = h(t) \langle \Theta, v \rangle
\end{aligned} \tag{5.10}(a)$$

for any $v \in V$ and with initial condition

$$w(0) = \dot{w}(0) = 0. \tag{5.10}(b)$$

Based on estimates similar to (5.7), it is straightforward to obtain the following result.

Lemma 5.2. Let (4.1) hold. Then there exists a unique solution w to problem (5.9).

Proof (of Proposition 5.1). From (5.8) it is clear that the first four terms of (5.1) converge to the first four terms of (5.9)(a). For the boundary terms we use the compactness of [10] to see that

$$w_{\delta_i} \rightarrow w \quad \text{in } L^2(0, T; L^q(\Gamma_C))$$

and

$$\dot{w}_{\delta_i} \rightarrow \dot{w} \quad \text{in } L^2(0, T; L^q(\Gamma_C))$$

Furthermore, from (5.7) it follows that

$$u(\beta + \delta h) \rightarrow u(\beta) \quad \text{in } H^1(0, T; V)$$

and therefore

$$u(\beta + \delta h) \rightarrow u(\beta) \quad \text{in } L^\infty(0, T; L^q(\Gamma_C))$$

where $1 \leq q \leq 4$ for $N=3$ and $1 \leq q$ for $N=2$.

Having observed the above convergence properties, we note that the boundary integral terms convey by the dominated convergence theorem.

Remark 5.1. We point out that the Sobolev regularization is used to obtain inequality (5.8). It is necessary since the coefficient of the $\|\dot{w}(t)\|_{V^2}^2$ term in inequality (5.7) cannot be made small.

Using the results from Proposition 5.1, we may now calculate the variation of J .

Corollary 5.3. Let $\text{meas}(\Gamma_D) > 0$, (3.2), (3.5), (3.6), (4.1) and (5.4) hold and let J^ε be given by (4.4)(b). Then $\partial J^\varepsilon(\beta)(h)$ is given by

$$\partial J^\varepsilon(\beta)(h) = 2 \int_0^T (u_\varepsilon(t; \beta) - z(t), w(t, h))_H dt + 2 \gamma(\beta, h)_{H^1(0, T)} \quad (5.11)$$

where w is the solution of (5.10).

Having determined the differentiability of the functional J^ε , we have the result.

Proposition 5.1. Let $\text{meas}(\Gamma_D) > 0$, (3.2), (3.5), (3.6), (4.1) and (5.4) hold. If β_ε is a solution of (4.4), then

$$\int_0^T (u_\varepsilon(t; \beta) - z(t), w(t, h))_H dt + \gamma(\beta_\varepsilon, h)_{H^1(0, T)} = 0 \quad (5.12)$$

for all $h \in L^2(0, T)$.

We may now use Proposition 5.1 to obtain regularity of β_ε , $\varepsilon > 0$.

Theorem 5.1. Let $\text{meas}(\Gamma_D) > 0$, (3.2), (3.5), (3.6), (4.1) and (5.4) hold. If β_ε is a solution of (4.4), then $\beta_\varepsilon \in H^2(0, T)$.

Proof. Consider the linear mapping

$$h \rightarrow \alpha(h) = \int_0^T (u_\varepsilon(t; \beta_\varepsilon) - z(t), w(t, h))_H dt$$

of $H^1(0, T)$ into R . We note from Proposition 4.2 and from (5.8) that

$$|\alpha(h)| \leq \|u_\varepsilon(\cdot, \beta_\varepsilon) - z(\cdot)\|_{L^2(0,T;H)} C \|h\|_{L^2(0,T)}$$

and thus in a continuous linear function on $L^2(0,T)$. It follows by the Riesz Representation theorem [16] that there exists $\tilde{\alpha} \in L^2(0,T)$ such that

$$\alpha(h) = (\tilde{\alpha}, h)_{L^2(0,T)}.$$

From (5.12) we see that for every $h \in H^1(0,T)$

$$(\beta_\varepsilon, h)_{H^1(0,T)} = -1/\gamma (\tilde{\alpha}, h). \quad (5.13)$$

We conclude from equation (5.13) that in fact

$$-\beta_{\text{ext}} + \beta_\varepsilon = -1/\gamma \tilde{\alpha}$$

with

$$\beta_\varepsilon(0) = \beta_\varepsilon(T) = 0.$$

Hence, $\beta_\varepsilon \in H^2(0,T)$.

6. APPROXIMATION

We begin by giving the Galerkin formulation for problem (4.3). Thus, let $S^N \subset V$ with basis $\{B_i\}_{i=1}^N$ and let

$$u^N(x,t) = \sum_{i=1}^N \mu_i(t) B_i(x)$$

be such that for $\varepsilon \geq 0$ fixed and $v \in S^N$

$$\begin{aligned} & \langle \ddot{u}_\varepsilon^N(t), v \rangle + \varepsilon a(\ddot{u}_\varepsilon^N(t), v) + c(u_\varepsilon^N(t), v) + a(\dot{u}_\varepsilon^N(t), v) + \langle P(u_\varepsilon^N(t), v) + \\ & + (J_\varepsilon(u_\varepsilon(t), \dot{u}_\varepsilon(t) - \Phi(t)), v) = \beta(t) \langle \Theta, v \rangle \end{aligned} \quad (6.1)(a)$$

with initial conditions

$$\begin{aligned} u_\varepsilon^N(0) &= u_0^N \\ \dot{u}_\varepsilon^N(0) &= u_1^N \end{aligned} \quad (6.1)(b)$$

where u_0^N and u_1^N converge to u_0 and u_1 in V , respectively, as $N \rightarrow \infty$. The following is not difficult to show.

Proposition 6.1. Let $\text{meas}(\Gamma_D) > 0$, (3.2), (3.5), (3.6), (4.1) and (5.4) hold and let $\varepsilon > 0$ be fixed. Let $\{\beta_i\}_{i=1}^\infty$ be a sequence in $H'(0, T)$ such that $\beta_i \rightarrow \beta$ weakly in $H'(0, T)$ and let $N_i \rightarrow \infty$.

Then

$$u_\varepsilon^{N_i}(\beta_i) \rightarrow u_\varepsilon(\beta) \text{ in } L^2(0, T; H)$$

as $i \rightarrow \infty$.

To approximate the controls, let L^M be a subspace in $H^1(0, T)$. Suppose there is a continuous linear mapping $I^M: H^1(0, T) \rightarrow L^M$ such that

$$\|u - I^M u\|_{H^1(0, T)} \leq C \|u\|_{H^1(0, T)}$$

and for $u \in H^2(0, T)$

$$\|u - I^M u\|_{H^1(0, T)} \leq \delta(M) \|u\|_{H^1(0, T)}$$

where $C \geq 0$ and $\delta(M) \geq 0$ with $\delta(M) \rightarrow 0$ as $M \rightarrow \infty$. Both C and $\delta(M)$ are independent of u .

The problems we consider are given by the following:

Find $\beta_2^{M,N} \in L^M$ such that

$$J_M^\varepsilon(\beta_2^{M,N}) = \inf \{ J_M^\varepsilon(\beta) : \beta \in L^N \} \quad (6.3)(a)$$

where

$$J_M^\varepsilon(\beta) = \int_0^T \| u^M(t; \beta) - z(t) \|_H^2 dt + \gamma \| \beta \|_{H^1(0,T)}^2 \quad (6.3)(b)$$

Remark 6.1. The existence of solutions $\beta_\varepsilon^{M,N}$ of (6.3). Moreover, these solutions are bounded in $H^1(0,T)$.

Theorem 6.1. Let $\tilde{\beta}_\varepsilon$ be a weak $H^1(0,T)$ limit point of $\beta_\varepsilon^{M_i, N_i}$ for $M_i \rightarrow \infty$ and $N_i \rightarrow \infty$. Then $\tilde{\beta}_\varepsilon$ is a solution of (4.4).

Proof. Since $\beta_\varepsilon^{M_i, N_i} \rightarrow \beta_\varepsilon$ weakly in $H^1(0,T)$ as $i \rightarrow \infty$, it follows from Proposition 6.1 that

$$u_\varepsilon^{M_i, N_i}(\beta_\varepsilon^{M_i, N_i}) \rightarrow u_\varepsilon(\tilde{\beta}_\varepsilon)$$

in $L^2(0,T)$ as $i \rightarrow \infty$. Thus, we see that

$$\lim_{i \rightarrow \infty} J_\varepsilon^{N_i, M_i}(\beta_\varepsilon^{M_i, N_i}) \geq J_\varepsilon(\tilde{\beta}_\varepsilon)$$

Now from Theorem 5.1, $\beta_\varepsilon \in H^1(0,T)$ and, thus from (6.2), we see that

$$J_\varepsilon^{N_i, M_i}(\beta_\varepsilon) \geq J_\varepsilon^{N_i, M_i}(\beta_\varepsilon^{M_i, N_i})$$

and

$$J_{\varepsilon}(\beta_{\varepsilon}) = \lim_{i \rightarrow \infty} J_{\varepsilon}^{N_i M_i N_i}(\beta_{\varepsilon}) \geq J_{\varepsilon}(\tilde{\beta}_{\varepsilon}) .$$

We conclude that $\tilde{\beta}_{\varepsilon}$ is a solution of (4.4).

Remark 6.1. It is now clear that an iterated limit theorem follows from Theorem 4.2 and Theorem 6.1.

Theorem 6.2. Let $\varepsilon_i \rightarrow 0$. Then there exist sequences $M_i N_i \rightarrow (\infty, \infty)$ as $i \rightarrow \infty$ such that

$$\text{w-lim}_{\varepsilon_i \rightarrow 0} (\text{w-lim}_{i \rightarrow \infty} \beta_{\varepsilon_j}^{M_i N_i}) = \beta_0$$

where β_0 is a solution of (3.16) and where the w-limits are weak limits in $H'(0,T)$. Also,

$$\lim_{i \rightarrow \infty} \lim_{j \rightarrow \infty} J_{\varepsilon_j}(\beta_{\varepsilon_j}^{M_i N_i}) = J(\beta_0) .$$

CHAPTER 6

QUALITATIVE AND NUMERICAL ANALYSIS OF ELASTOHYDRODYNAMIC LUBRICATION PROBLEMS

1. INTRODUCTION

This work concerns the mathematical and numerical analysis of elastohydrodynamic lubrication problem for both one-dimensional line contact and two-dimensional point contact cases.

Lubrication has been a very active research field for several decades. Because of its importance in machine design and general mechanical engineering, for nearly a century, it is a subject that has attracted the concern and energy of numerous researchers and experimentalists in diverse industries throughout the world. Dowson has written a voluminous treatise [27] on the long history and the development of lubrication theory and practice. Among the many aspects of lubrication engineering, elastohydrodynamic lubrication (EHL) has received considerable interest in recent years, and methods for overcoming the complexity and strong nonlinearity of EHL problems have been the subject of numerous articles, several Journals, and a series of world conferences.

Starting from Grubin [37] and Petrusевич [70]'s efforts in late 1940's and early 1950's in integrating the Reynolds equation for a simplified two-dimensional model, many subsequent papers were devoted to further studies of this model. Especially in the past decade, modern computers enabled analysts to obtain numerical simulations of certain classes of lubrication problems.

Despite significant advances in EHL theory, the inherent nonlinearity of the governing equations has stood in the way of a fully satisfactory study, either mathematical or numerical, of wide classes of problems of interest to engineers. Very few papers can be found on the analysis for the mathematical behavior of the solutions or the numerical accuracy of discrete solutions. Also because of the difficulty of obtaining accurate experimental results, reliable numerical analysis has become essentially important to today's tribologists.

The objective of this report is fourfold. First, to develop a new family of finite element methods for treating the free boundary problems of

elastohydrodynamic lubrication characterized by the equations and inequalities of the Reynolds–Hertz theory. To handle the problem of determining the free boundary where the contact pressure goes to zero, we introduce an exterior penalty method. Second, to study mathematical features of the problem, including the characterization of governing equations, the existence of solutions, the convergence of penalty approximations, the appropriateness of certain boundary conditions, and very importantly, the regularity of solutions to the general line–contact problem. Thirdly, to construct numerical solutions to representative line–contact and point–contact problems in lubrication theory and, finally, to develop a–priori error estimates and to prove convergence of finite element approximations of certain classes of elastohydrodynamic lubrication problems.

Following this introduction, the governing equations for elastohydrodynamic lubrication, as described by the simplified Reynolds–Hertz theory, are derived and summarized.

In Section 3, a variational formulation of the free boundary problem of lubrication is derived which is characterized by a nonlinear variational inequality. In this chapter, a penalty method is also introduced to regularize the problem by releasing the constraint on the cavitation boundary.

In Section 4, properties of the governing operator in the variational problem are studied. It is proved that the operator is bounded, coercive, pseudomonotone, and continuous. These properties are then used to prove a basic existence theorem for the solutions. Further study reveals the convergence of solutions to the penalized problem to a solution of the variational inequality as the penalty parameter ϵ tends to zero. Also in this chapter, the regularity of solutions to the one–dimensional line contact problems is studied and regularity theorems are established.

Section 5 presents a finite element formulation of the penalized problem. Numerical results are presented which confirm the theoretical analysis. The penalty method is shown to be effective in locating the free boundary.

Section 6 is devoted to the study of properties of approximations to the lubrication problem in finite dimensional spaces which is generally the foundation for numerical analysis. The study establishes convergence criteria and relations

among the original problem, the penalized problem, and their finite element approximations. An a-priori error estimate for finite element approximations of a "light load" case is presented. Numerical experiments show good agreement with the predicted rates of convergence for several test cases. Then the adaptive-refinement method is applied to the elastohydrodynamic lubrication problems.

Section 7 lists several open problems for future study, and summarizes basic accomplishments and conclusions of the study.

2. FUNDAMENTAL EQUATIONS

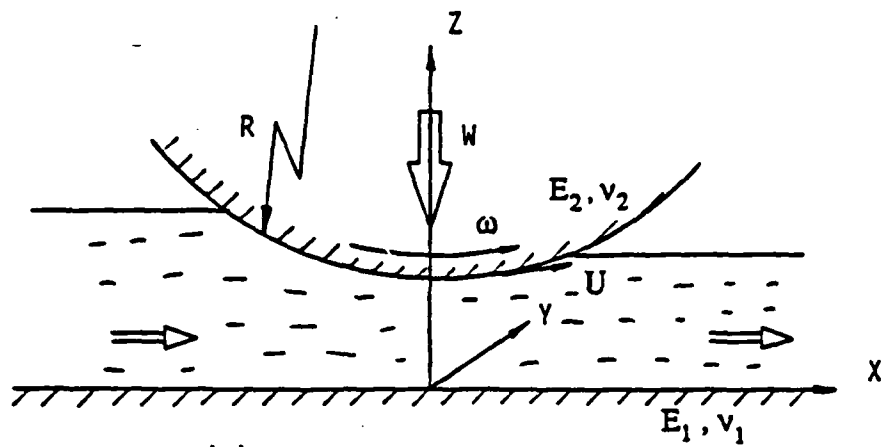
A full century has elapsed since Osborne Reynolds [75] derived equations to simulate the pressure on lubricated surfaces and thus began the mathematical theory of lubrication.

Guided by experimental results of Beauchamp Tower [89] and Nikolai P. Petrov [69] which established the existence of a thin film of lubricant between the lubricated solid machine parts, Reynolds developed a reduced version of the Navier-Stokes equations to model the lubricant. The resulting model of lubricated contact has been extremely successful in engineering lubrication problems and is now standard fare in virtually all textbooks on tribology, e.g., [9,28,42,71].

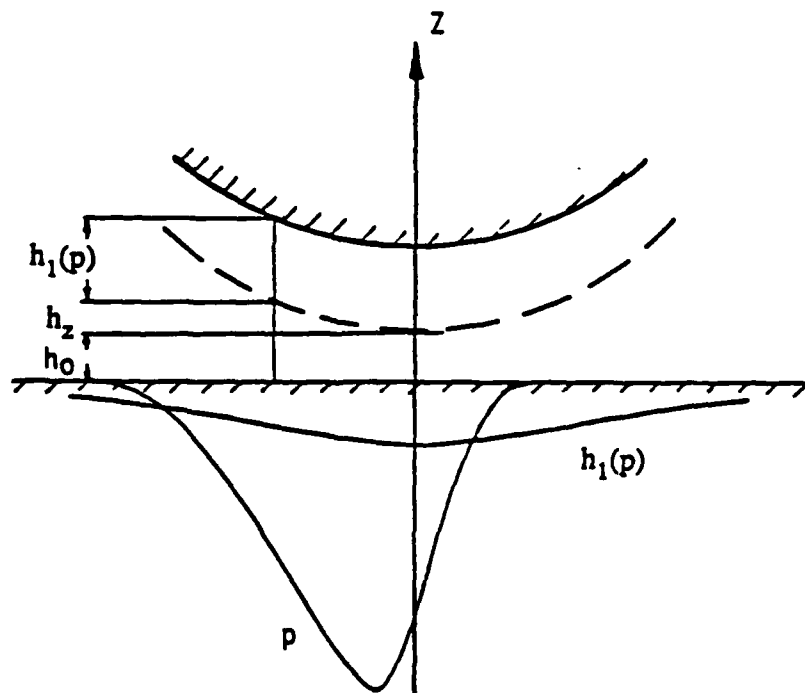
Here we record the derivation of a generalization of Reynolds equations which take into account small deformations of the solids. The results of Hertz on elastic contact are applicable to this class of problems, and the theory is sometimes referred to as the Reynolds-Hertz theory of elastohydrodynamic lubrication.

2.1 MODEL.

The physical situation of interest here is the behavior of a cylindrical or spherical elastic roller being pushed onto an essentially flat interface, with the roller and interface being separated by a thin film of lubricant. Figure 2.1 shows a cylindrical or spherical elastic roller of radius R carrying load W . The roller is rotating about an axis parallel to y axis in the figure, and the flat interface is modeled as a half space. A thin film of lubricant separates the bodies and flows through the gap between them, the applied load on the roller being balanced by the pressure developed in the lubricant. Both the roller and the support suffer deformations due to the lubricant pressure and this elastic deformation, in turn, alters the geometries of the lubricated region between them. The problem is thus one of fluid-solid interaction, as the flow field of the lubricant is altered by the deformation and this results in a redistribution of fluid pressure.



(a) Mechanism of Bearing



(b) Pressure and Deformation

Fig. 2.1 Mechanism of Bearing

Starting with a general case, we denote by U_1, V_1, W_1 and U_2, V_2, W_2 the velocity components of the two solid surfaces, as the boundary values of lubricant velocities; h denotes the actual thickness of the gap, R the radius of the roller and ω the angular velocity. The pressure, density and viscosity of the lubricant are denoted p , ρ , and μ , respectively, and E_1, ν_1 and E_2, ν_2 denote the Young's moduli and Poisson's ratios for the respective solids. The lubricant flows along x -direction in the figure and is assumed to have an ambient pressure $p = 0$ at both the inlet and the outlet.

Customally, the case of a cylindrical roller can be studied as a one-dimensional problem of line contact and the case of a spherical roller can be treated as a two-dimensional point contact problem. We establish a Cartesian coordinate system with the x, y - axes in the plane of the solid interface and z normal to the plane, as shown in Fig. 2.1a.

2.2 SIMPLIFYING ASSUMPTIONS.

Various models of lubrication may differ according to the definition of the material properties of lubricants, the geometry of bearing configurations (thrust, slider, roller, etc.), the assumed effects of surface roughness, temperature, dynamic loadings, etc. Here we confine ourselves to a typical situation in which the following assumptions are enforced:

- (1) The lubricant is characterized as a non-Newtonian fluid with a pressure-dependent viscosity. In particular, the relation between deviatoric stress and velocity gradients is of the form

$$\tau_{ij} = \mu/2 (\partial v_i / \partial x_j + \partial v_j / \partial x_i)$$

where $\mu = \mu(p) = \mu_0 e^{-\alpha p}$ and p is the "hydrostatic" fluid pressure.

- (2) The flow of the lubricant is laminar.
- (3) No cavitation or slip occurs within the contact region.
- (4) The physical state is isothermal.
- (5) The effects of body and inertia forces are negligible.
- (6) The roller and bearing are linearly elastic with smooth surfaces. (H-1)

The Navier-Stokes equation for the Newtonian fluid is

$$\rho Du_i/Dt = \rho f_i - p_{,i} - 2/3(\mu u_{k,k})_{,i} + (\mu(u_{i,j} + u_{j,i}))_{,j} \quad (2.1)$$

The continuity equation is

$$\frac{\partial \rho}{\partial t} + (\rho u_i)_{,i} = 0 \quad (2.2)$$

where the conventional tensor notation of repeated indices are used for summation from 1 to 3; u_1, u_2 and u_3 denote the velocity components u, v and w respectively, x_1, x_2 and x_3 for x, y and z , and ρf_i express the body forces.

The thinness of the film of lubricant suggests that additional assumptions can be introduced:

- (a) The film thickness h is "small", i.e., h is much smaller than the characteristic length of the contact region and much smaller than the radius of curvature of the contact surface.
- (b) Variations of pressure, density and viscosity in thickness direction are much smaller than those in the other directions. Thus we assume

$$p = p(x, y), \quad \rho = \rho(x, y), \quad \mu = \mu(x, y) \quad (2.3)$$

- (c) The velocity gradient is greater in thickness direction than in the other directions. That means

$$\partial u_i / \partial x, \partial u_i / \partial y \ll \partial u_i / \partial z \quad (H-2)$$

2.3 REYNOLDS EQUATION.

From the above hypotheses (H-1) and (H-2), after neglecting the inertia and body forces, and the small terms of the velocity gradients, the Navier-Stokes equations are reduced to

$$\begin{cases} \frac{\partial p}{\partial x} = \frac{\partial}{\partial z} \left(\mu \frac{\partial u}{\partial z} \right) = \mu \frac{\partial^2 u}{\partial z^2} \\ \frac{\partial p}{\partial y} = \frac{\partial}{\partial z} \left(\mu \frac{\partial v}{\partial z} \right) = \mu \frac{\partial^2 v}{\partial z^2} \end{cases} \quad (2.4)$$

Integrating (2.4) with respect to z , we have

$$\begin{cases} u = \frac{\partial p}{2\mu\partial z} z^2 + A_1 z + A_0 \\ v = \frac{\partial p}{2\mu\partial z} z^2 + B_1 z + B_0 \end{cases} \quad (2.5)$$

Using the boundary conditions at solid surfaces:

$$\begin{cases} u = U_2, v = V_2, w = W_2 & \text{at } z = h \\ u = U_1, v = V_1, w = W_1 & \text{at } z = 0 \end{cases} \quad (2.6)$$

we have

$$\begin{cases} u = \frac{\partial p}{2\mu\partial x} z(z-h) + \frac{h-z}{h} U_1 + \frac{z}{h} U_2 \\ v = \frac{\partial p}{2\mu\partial y} z(z-h) + \frac{h-z}{h} V_1 + \frac{z}{h} V_2 \end{cases} \quad (2.7)$$

Substituting (2.7) into continuity equation (2.2) and using the identity

$$\frac{\partial}{\partial \xi} \int_0^{h(\xi)} f(\xi, \zeta) d\zeta = \int_0^{h(\xi)} \frac{\partial f}{\partial \xi} d\zeta + f(\xi, h(\xi)) \frac{\partial h}{\partial \xi}$$

we have

$$\begin{aligned} & h \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} \left(-\frac{\rho}{12\mu} h^3 \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(-\frac{\rho}{12\mu} h^3 \frac{\partial p}{\partial y} \right) + \rho(W_2 - W_1) \\ & + \frac{\partial}{\partial x} \left(\frac{U_1 + U_2}{2} \rho h \right) + \frac{\partial}{\partial y} \left(\frac{V_1 + V_2}{2} \rho h \right) - \rho U_2 \frac{\partial h}{\partial x} - \rho V_2 \frac{\partial h}{\partial y} = 0 \end{aligned} \quad (2.8)$$

From the geometry of the roller surface (see Fig. 2.2), we have

$$W_2 - W_1 = \frac{\partial h}{\partial x} U_2 + \frac{\partial h}{\partial y} V_2 + \frac{\partial h}{\partial t} \quad (2.9)$$

(2.8) is hence simplified as

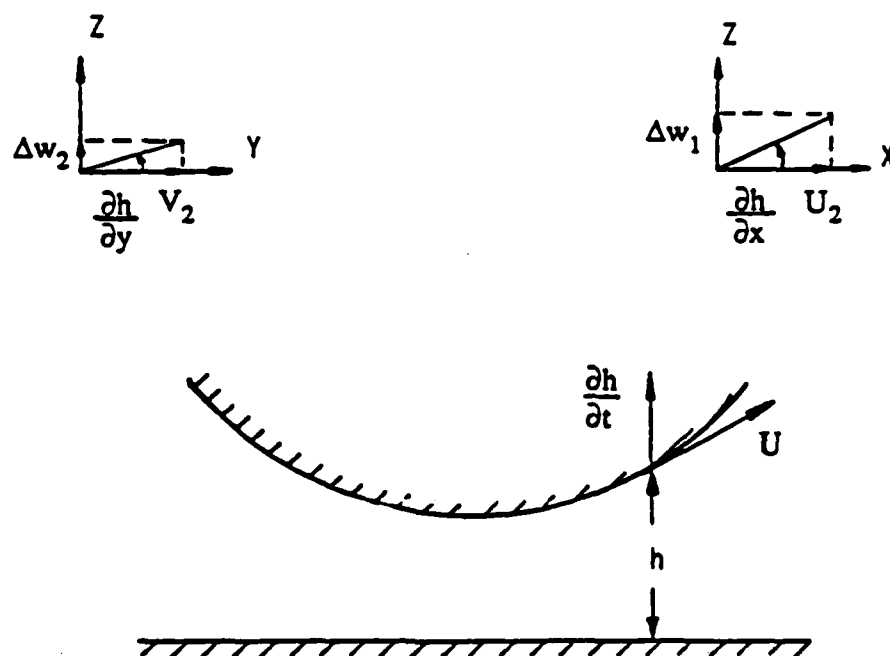


Fig. 2.2 Velocity at Roller Surface

$$\begin{aligned}
& -\frac{\partial}{\partial x} \left(\frac{\rho}{12\mu} h^3 \frac{\partial p}{\partial x} \right) - \frac{\partial}{\partial y} \left(\frac{\rho}{12\mu} h^3 \frac{\partial p}{\partial y} \right) + \frac{\partial(\rho h)}{\partial t} \\
& + \frac{\partial}{\partial x} \left(\frac{U_1 + U_2}{2} \rho h \right) + \frac{\partial}{\partial y} \left(\frac{V_1 + V_2}{2} \rho h \right) = 0
\end{aligned} \tag{2.10}$$

Suppose, for a typical case, there is no leakage, so $V_1 = V_2 = 0$, and denote $u_0 = (U_1 + U_2)/2$. Then (2.10) becomes the classical Reynolds equation

$$-\frac{\partial}{\partial x} \left(\frac{\rho}{12\mu} h^3 \frac{\partial p}{\partial x} \right) - \frac{\partial}{\partial y} \left(\frac{\rho}{12\mu} h^3 \frac{\partial p}{\partial y} \right) + \frac{\partial}{\partial x} (u_0 \rho h) + \frac{\partial}{\partial t} (\rho h) = 0 \tag{2.11}$$

At this point, $h = h(x, y)$ is regarded as a given function of (x, y) . Of course, it becomes a function of p when the deformation of bearing is taken into account.

2.4 ELASTICITY.

The elastic deformation of the surface of a linearly elastic half space subject to distributed loads normal to the surface can be found in the classical elasticity texts. Referring to [50], we have for cylindrical contact (plane strain model), Michell's solution

$$w(x) = \frac{1-\nu^2}{\pi E} \int p(\xi) \ln \left(\frac{x_0 - \xi}{x - \xi} \right)^2 d\xi \tag{2.12}$$

Here w is the normal displacement of a point x on the surface of the half space and x_0 is a remote point where no deformation is significant; and for spherical contact, the solution of Boussinesq's problem is applicable:

$$w(x, y) = \frac{1-\nu^2}{\pi E} \int \frac{p(\xi, \eta)}{\sqrt{(x-\xi)^2 + (y-\eta)^2}} d\xi d\eta \tag{2.13}$$

When we consider elastic contact between two solids, the total deformation should be the summation of contributions of the two parts. Then we may modify (2.12) and (2.13) as

$$h_1(p) = \frac{2}{\pi E'} \int_{\Omega} \ln (x_0 - \xi / x - \xi)^2 p d\xi \quad \text{for 1-D contact} \tag{2.14}$$

$$h_1(p) = \frac{2}{\pi E'} \int_{\Omega} \frac{1}{r} p \, d\xi \, d\eta \quad \text{for 2-D contact} \quad (2.15)$$

where r is the distance between points x and ξ , and E' is the effective elastic modulus, defined as

$$\frac{1}{E'} = \frac{1}{2} \left(\frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \right) \quad (2.16)$$

In classical elastohydrodynamic theory, it is assumed that the roller and bearing deform under the lubricant pressure as if they were in direct contact. Thus (2.14–16) can be used to evaluate the elastic deformation of the domain surrounding the lubricant.

Henceforth, we express the film thickness, including the elastic deformation, as

$$h(p) = h_1(p) + h_2$$

where

$$\begin{cases} h_2 = h_0 + h_z \\ h_z = R - \sqrt{R^2 - x^2} \end{cases} \quad \text{for 1-D contact}$$

$$R - \sqrt{R^2 - x^2 - y^2} \quad \text{for 2-D contact} \quad (2.17)$$

h_0 is a reference parameter indicated in Fig. 2.1(b). If the roller could recover its deformation in loaded configurations with its center unmoved, then the least distance would be h_0 .

Technically we may also control the minimum thickness h_m , shown in Fig. 2.3, instead of h_0 in computation. Then

$$\begin{cases} h(p) = \tilde{h}_1 + \tilde{h}_2 \\ \tilde{h}_2 = h_z + h_m, \quad h_m = \text{const.} \\ \tilde{h}_1 = h_1(p) - s(h), \quad s(h) = \inf (h_1(p) + h_z) \end{cases} \quad (2.17')$$

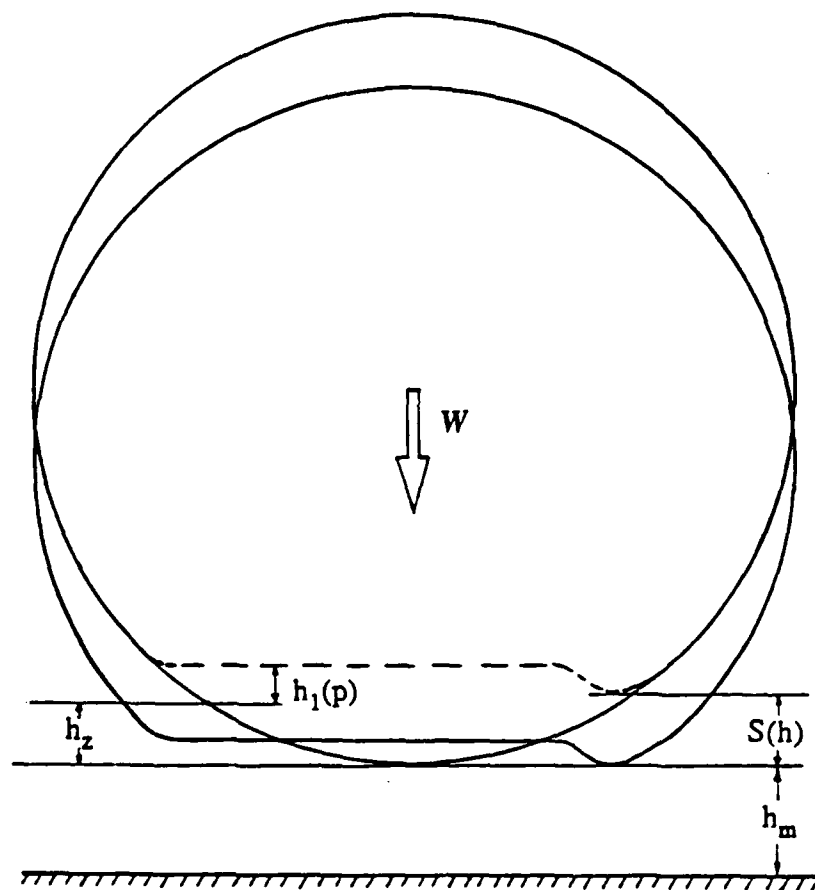


Fig. 2.3 The Minimum Film Thickness

We are to consider the simplified case with a rigid half space ($E_1 = \infty$). Thus (2.14–15) reduce to (2.12–13). Due to the dependence of h and μ on p , this expanded version of Reynolds theory (with (2.14) or (2.15)) is now characterized by a kind of differential–integral equation, refer to this as the Reynolds–Hertz equation, because of the use it makes with Hertz type contact behavior.

2.5 BOUNDARY CONDITION. CAVITATION-FREE BOUNDARY.

We shall consider fully lubricated bearings. That means, there can be no cavitation occurring inside the contact region where $p > 0$. A lubricant with ambient pressure enters the contact region and leaves there with a recovered ambient pressure. Thus if a large enough domain Ω is selected wherein includes the entire contact region Ω_1 , we must then have

$$p = 0 \quad \text{at } \partial\Omega, \text{ the boundary of } \Omega \quad (2.18)$$

$$p = 0 \quad \text{at } \partial\Omega_1, \text{ the boundary of } \Omega_1 \quad (2.19)$$

and in $\Omega_0 = \Omega - \Omega_1$, $p = 0$.

The location of Ω_1 can not be determined a-priori. Like most contact problems, the Reynolds–Hertz problem is a type of free boundary problem.

On the other hand, the flux per unit contact length can be obtained by integrating the momentum ρu and ρv along the thickness. By virtue of (2.7)

$$\begin{cases} q_x = \int_0^h \rho u \, dz = -\frac{\rho}{12\mu} \frac{\partial p}{\partial x} h^3 + \frac{U_1 + U_2}{2} \rho h \\ q_y = \int_0^h \rho v \, dz = -\frac{\rho}{12\mu} \frac{\partial p}{\partial y} h^3 \end{cases} \quad (2.20)$$

The boundary condition can also be determined by specifying a given flux. However these fluxes are seldom known, we shall focus here on Dirichlet type conditions.

Since outside the contact region we always have $p = 0$, we also have there

$\partial p/\partial x=0$, $\partial p/\partial y=0$ if the pressure is smooth enough. From the continuity of flow, Dowson and Higginson^[28] assert the so-called Reynolds condition for one-dimensional case

$$\left. \frac{dp}{dx} \right|_{\partial\Omega_1} = 0 \quad (2.21)$$

For two-dimensional case, this condition assumes the form

$$\nabla p|_{\partial\Omega_1} = 0 \quad (2.22)$$

Some investigators have used these conditions as the criteria to locate $\partial\Omega_1$ in the numerical procedures. However, these conditions can hold only if the solutions are sufficiently smooth.

2.6 SUMMARY.

In summary, for constant density ρ and a modified Newtonian fluid with $\mu=\mu_0 e^{-\alpha p}$, the flow of a lubricant through elastic bearing is described by the Reynolds-Hertz equations and boundary conditions:

$$\begin{cases} -\nabla \cdot (h^3(p) e^{-\alpha p} \nabla p) + 12\mu_0 \frac{\partial(uh(p))}{\partial x} + 12\mu_0 \frac{\partial h(p)}{\partial t} = 0 \\ p \geq 0 \end{cases} \quad \text{in } \Omega_1 \quad (2.23)$$

$$p = 0 \quad \text{in } \Omega_0 \quad (2.24)$$

$$p|_{\partial\Omega} = 0 \quad \Omega = \Omega_0 \cup \Omega_1 \quad \text{with undetermined } \partial\Omega_1 \quad (2.25)$$

Here the operator ∇ denotes d/dx for one-dimensional problems and $(\partial/\partial x, \partial/\partial y)$ for two-dimensional problems.

Conditions (2.24) and (2.25) may be replaced by

$$p|_{\partial\Omega_1} = 0 \quad \nabla p|_{\partial\Omega_1} = 0 \quad \text{with undetermined } \partial\Omega_1 \quad (2.26)$$

In (2.23),

$$h = \tilde{h}_1 + \tilde{h}_2 \quad (2.27)$$

$$\tilde{h}_1 = h_1(p), \quad h_2 = h_0 + h_z \quad (2.28)$$

or

$$\tilde{h}_1 = h_1(p) - s(h), \quad \tilde{h}_2 = h_m + h_z, \quad s(h) = \inf (h_z + h_1(p)) \quad (2.28')$$

In the line contact case

$$\begin{cases} h_z = R - \sqrt{R^2 - x^2} \\ h_1(p) = \frac{2}{\pi E'} \int p(\xi) \left(\frac{x_0 - \xi}{x - \xi} \right)^2 d\xi, \quad x_0 \text{ is a remote point} \end{cases} \quad (2.29)$$

In the point contact case

$$\begin{cases} h_z = R - \sqrt{R^2 - x^2 - y^2} \\ h_1(p) = \frac{2}{\pi E'} \iint \frac{p(\xi, \eta)}{\sqrt{(x - \xi)^2 + (y - \eta)^2}} d\xi d\eta \end{cases} \quad (2.30)$$

3. VARIATIONAL FORMULATION — PENALTY METHOD

The highly nonlinear character of the equations and inequalities developed in the preceding chapter make it necessary to resort to numerical methods to obtain solutions to the Reynolds–Hertz theory. In the present chapter, we develop a penalty formulation of the problem as a basis for numerical schemes to be derived later.

We start with the definition of a nonlinear operator for the steady–state problem. Formally, we can define an operator

$$A(p) \equiv -\nabla \cdot (h^3(p)e^{-\alpha p} \nabla p) + 12\mu_0 \frac{\partial u_0 \tilde{h}_1(p)}{\partial x} \quad (3.1)$$

and a function

$$f \equiv -12\mu_0 \frac{\partial (u_0 \tilde{h}_2)}{\partial x} \quad (3.2)$$

We shall define the domain and codomain for A later.

Then we write the governing equations as:

$$\begin{cases} A(p) = f, & p > 0 & \text{in } \Omega_1 \\ p = 0 & & \text{in } \Omega_0 \\ p|_{\partial\Omega} = 0, & \Omega = \Omega_0 \cup \Omega_1 \end{cases} \quad (3.3)$$

with (2.27–30) for evaluating $h(p)$.

Free boundary–value problems of the type (3.3) can often be characterized as variational inequalities, particularly if the condition defining the free boundary ($p \geq 0$) defines a closed convex set in the domain of the governing operator. We shall now show that the Reynolds–Hertz problem can, indeed, be formulated as a variational inequality. This fact is significant, not only because it makes available a formidable arsenal of mathematical results on the qualitative behavior of the

solutions of variational inequalities, but also because the formulations and theory suggest effective methods for obtaining numerical solutions.

3.1 PRELIMINARIES ON VARIATIONAL INEQUALITIES.

We are now interested in a detailed mathematical analysis of (3.3). We begin by reviewing a number of mathematical definitions and concepts relevant to the theory of variational inequalities. In particular, we cite several basic definitions and important theorems, proofs of which can be found in Oden and Kikuchi^[57] or [30,33,35,55]. For additional detail, see the monograph of Oden^[55]. The notations are in consistence with the classical texts, e.g., [54,56,79].

We denote by U' for the topological dual of a normed linear space U with the norm $\|\cdot\|_U$. The canonical duality pairing on $U' \times U$ is denoted

$$\langle \cdot, \cdot \rangle : U' \times U \rightarrow \mathbb{R} \text{ (real numbers)} \quad (3.4)$$

Thus for a linear functional on U , $f : U \rightarrow \mathbb{R}$, we write

$$f(u) = \langle f, u \rangle$$

DEFINITION 3.1. A sequence $\{u_n\}$ of U converges strongly to $u \in U$, iff

$$\lim_{n \rightarrow \infty} \|u_n - u\|_U = 0 \quad (3.5)$$

and $\{u_n\}$ converges weakly to $u \in U$ iff

$$\lim_{n \rightarrow \infty} \langle f, u_n \rangle = \langle f, u \rangle, \quad \forall f \in U' \quad (3.6)$$

DEFINITION 3.2. Let functional f be defined on a subset K of U .

If for any sequence $\{u_n\} \in K$ converging weakly to $u \in K$, we have

$$\liminf_{n \rightarrow \infty} f(u_n) \geq f(u) \quad (3.7)$$

f is said to be weakly (sequentially) lower semicontinuous at u .

DEFINITION 3.3. A subset K of U is convex iff

$$\theta x + (1-\theta)y \in K, \quad \forall x, y \in K, \text{ for } 0 \leq \theta \leq 1 \quad (3.8)$$

DEFINITION 3.4. Functional f defined on a convex set $K \subset U \rightarrow R$ is convex iff

$$f(\theta u + (1-\theta)v) \leq \theta f(u) + (1-\theta)f(v), \quad \forall u, v \in K, \theta \in [0,1] \quad (3.9)$$

DEFINITION 3.5. (1) An operator $A : K \subset U \rightarrow U'$ is monotone on K iff

$$\langle A(u) - A(v), u - v \rangle \geq 0 \quad \forall u, v \in K \quad (3.10)$$

If it attains 0 only when $u=v$, A is said to be strictly monotone.

(2) A is hemicontinuous at u , if the function $f : [0,1] \rightarrow R$ defined by

$$f(t) = \langle A(u + tv), w \rangle \quad (3.11)$$

is continuous for all $v, w \in U$.

(3) A is coercive on K if there exists $u_0 \in K$, for $u \in K$, such that

$$\frac{\langle A(u), u - u_0 \rangle}{\|u\|_U} \rightarrow +\infty \quad \text{as } \|u\|_U \rightarrow \infty \quad (3.12)$$

DEFINITION 3.6. A functional f defined on normed linear space V is called Gâteaux differentiable at u , if there exists a continuous linear functional $g \in V'$ such that, for any $v \in V$,

$$\lim (f(u + \theta v) - f(u)) / \theta = \langle g, v \rangle \quad (3.13)$$

Moreover, such a g is denoted $Df(u)$, and is called the Gâteaux differential of f at u .

DEFINITION 3.7. Let A denote an operator, $A : K \subset U \rightarrow U'$, defined on a nonempty closed convex set K in a real linear topological space U , and let f be given in U' . The problem of finding $u \in K$ such that

$$\langle A(u) - f, v - u \rangle \geq 0, \quad \forall v \in K \quad (3.14)$$

is a variational inequality for the operator A .

DEFINITION 3.8. Let V be a Banach space, K a non-empty, closed

convex subset of V . An operator $A : K \subset V \rightarrow V'$ and a sequence $\{u_n\} \in K$ have the property P if

[P] : $u_n \rightarrow u$ weakly in V and

$$\lim_{n \rightarrow \infty} \langle A(u_n), u_n - u \rangle \leq 0 \quad (3.15)$$

An operator A is pseudomonotone on K iff whenever A and $\{u_n\}$ satisfy property P, we have

$$\lim_{n \rightarrow \infty} \langle A(u_n), u_n - v \rangle \geq \langle A(u), u - v \rangle, \quad \forall v \in K \quad (3.16)$$

Listed below are the important results in the theory of variational inequalities.

Consider an abstract variational inequality on a Banach space. Let

U be a separable reflexive real Banach space.

$K \subset U$ be a non-empty closed convex subset of U .

$A : K \subset U \rightarrow U'$ be an operator defined on K . (3.17)

THEOREM 3.1. Under conditions (3.17), if A is (1) bounded, (2) pseudomonotone on K , and (3) if K is bounded, then there exists at least one solution u of the variational inequality

$$\langle A(u), v - u \rangle \geq 0, \quad \forall v \in K \quad (3.18)$$

THEOREM 3.2. Under conditions (3.17) and the conditions (1) and (2) in above Theorem 3.1 for operator A , and for any unbounded set K , (3.18) has at least one solution if A satisfies the coerciveness condition (3.12) or a weak coerciveness condition:

$\exists v_0 \in K$ and $r > 0$ with $\|v_0\|_U < r$ such that

$$\langle A(v), v - v_0 \rangle > 0, \quad \forall v \in K, \text{ with } \|v\|_U = r \quad (3.19)$$

THEOREM 3.3 Under the conditions in Theorem 3.2, if A satisfied the coercivity of (3.12), then for any $f \in U'$, there is solution

$$u \in K, \langle A(u) - f, v - u \rangle \geq 0, \quad \forall v \in K \quad (3.20)$$

Also we may apply these properties to an equation.

THEOREM 3.4. If A is bounded, coercive and pseudomonotone, then A is surjective, i.e., $\forall f \in U'$, there exists a solution $u \in U$ of the equation $A(u) = f$.

COROLLARY 3.5. Under the conditions of Theorem 3.1–3.4, if A is strictly monotone (which implies that A is pseudomonotone), then the solution is unique.

In Chapter 4, some conditions to guarantee an operator to be pseudomonotone will be introduced so that we can use the theorems listed here.

3.2 VARIATIONAL INEQUALITY — WEAK STATEMENT.

We may construct a variational statement for our problem in order to obtain the weak solution.

Define the operator A in (3.1) on the Sobolev space

$$V = H^1_0(\Omega) = \{v, \partial v / \partial x, \partial v / \partial y \in L^2(\Omega), v|_{\partial\Omega} = 0\} \quad (3.21)$$

for $f \in L^2(\Omega)$, the duality pairing on f is

$$\langle f, v \rangle = \int_{\Omega} f v \, d\Omega \quad (3.22)$$

Then A maps V into its topological dual space $V' = H^{-1}$, and f in (3.2) is in fact in $C^\infty(\Omega) \subset V'$. Define a subset K of V by

$$K = \{v \in V \mid v \geq 0 \text{ a.e. in } \Omega\} \quad (3.23)$$

Evidently K is non-empty, closed and convex in V .

We may thus construct a variational inequality for problem (3.3), referred to hereafter as problem (P):

(P): Find $p \in K$ such that

$$\langle A(p) - f, q - p \rangle \geq 0, \quad \forall q \in K \quad (3.24)$$

with A , the set K , and the space V defined by (3.1), (3.21) and (3.23) respectively.

That problem P is in some sense equivalent to our lubrication problem is established in the following result.

THEOREM 3.6. If p is a solution of the variational inequality (3.24), then p is a solution of (3.3) in distributional sense; i.e., $p \in H^1_0(\Omega)$, $p \geq 0$ in Ω , $A(p)-f = 0$ in distributional sense in $\Omega_1 \subset \Omega$ where $p > 0$. (Thus p is referred to as a weak solution, or simply an H^1 -solution.)

PROOF: For the solution p of (3.24), define

$$\Omega_1 = \{(x,y) \in \Omega \mid p(x,y) > 0\}; \quad \Omega_0 = \{(x,y) \in \Omega \mid p(x,y) = 0\}.$$

Let $\psi \in C^\infty_0(\Omega_1)$ be such that $\psi > 0$ in Ω_1 and extend ψ to Ω by defining $\psi = 0$ in Ω_0 .

Denote $q = p + \varepsilon\psi$, for small ε , $q \geq 0$, $q \in K$. Then (3.24) results in $\varepsilon \langle A(p)-f, \psi \rangle \geq 0$. Since ε can be set positive or negative, we must have $\langle A(p)-f, \psi \rangle = 0$. That means $A(p) - f = 0$ in Ω_1 in a distributional sense.

REMARK. Returning to (3.24), by virtue of the fact that $p=0$ in Ω_0 and $A(p)-f=0$ in Ω_1 (distributional sense), we have

$$\langle A(p)-f, q-p \rangle = \int_{\Omega} (A(p)-f) (q-p) d\Omega = \int_{\Omega_0} (A(p)-f) q d\Omega_0 \geq 0.$$

$q \in K$, $q \geq 0$ implies a necessary condition on the data:

$$A(p)-f \geq 0 \text{ in } \Omega_0 \tag{3.25}$$

In our case, $A(p) = 12\mu_0 \partial[u_0 \tilde{h}_1(p)]/\partial x$ in Ω_0 , $u_0 = 0.5\omega R \cos\theta$, $\cos\theta = \sqrt{R^2-x^2-y^2}/R$, $f = -12\mu_0 \partial(u_0 h_2)/\partial x$. So we need

$$\partial[\sqrt{R^2-x^2-y^2} (\tilde{h}_1(p) + \tilde{h}_2)]/\partial x = \partial(\sqrt{R^2-x^2-y^2} h)/\partial x \geq 0 \text{ in } \Omega_0.$$

$$\tilde{h}_2 = h_0 + h_z = h_0 + R(1 - \sqrt{R^2-x^2-y^2}/R), \quad \partial \tilde{h}_2/\partial x = x/\sqrt{R^2-x^2-y^2}$$

$$\begin{aligned}\tilde{h}_1(p) &= C_E \int_{\Omega} K(x, \xi) p(\xi) d\Omega = C_E \int_{\Omega_1} K(x, \xi) p(\xi) d\Omega \\ &= C_E \int \ln(x_0 - \xi / x - \xi)^2 p(\xi) d\xi \quad \text{line contact} \\ C_E \int p(\xi, \eta) / \sqrt{(x - \xi)^2 + (y - \eta)^2} d\Omega &\quad \text{point contact}\end{aligned}$$

$$\begin{aligned}\partial \tilde{h}_1(p) / \partial x &= \partial h_1(p) / \partial x = C_E \int_{\Omega_1} -2p(\xi) / (x - \xi) d\xi \quad \text{line contact} \\ C_E \int_{\Omega_1} -(x - \xi) / r^3 p(\xi, \eta) d\Omega &\quad \text{point contact}\end{aligned}$$

The outlet zone Ω_0 should be in the down stream of the contact region. For line contact, if $\xi \in \Omega_1$ and $x \in \Omega_0$, we have $x \geq \xi$. So we always have $\partial \tilde{h}_1(p) / \partial x \leq 0$. With the hypotheses (H-2a), (3.25) gives

$$\begin{aligned}0 &\leq -xh / \sqrt{R^2 - x^2 - y^2} + \sqrt{R^2 - x^2 - y^2} (\partial h_1(p) / \partial x + x / \sqrt{R^2 - x^2 - y^2}) \\ &\leq x (1 - h / \sqrt{R^2 - x^2 - y^2}) < x.\end{aligned} \quad (3.25')$$

This indicates that the domain Ω_0 is located down stream, and the flow is possible.

For point contact, consider point $M \in \Omega_1$ where $\xi_M = \sup \{ \xi \mid (\xi, \eta) \in \Omega_1 \}$.

For points $(x, \eta_M) \in \Omega_0$, $x > \xi$, we still have $\partial h_1(p) / \partial x \leq 0$. And (3.25') still holds.

Condition (3.25) is thus physically accessible. Numerical solutions confirm the condition (3.25'), see e.g., [28,42,93].

3.3 PENALIZED FORMULATION.

The penalty method has been utilized extensively, particularly in dealing with the contact problems. Oden and Kikuchi^[57], and Pires^[72] successfully developed formulations, analyses, and numerical schemes for certain kinds of free boundary problems. As an application to our problem, we introduce a penalty term,

p^-/ϵ , defined by

$$p^- = \min(p, 0) = (p - |p|) / 2 \quad (3.26)$$

When regularizing the constraint, we consider the penalized problem:

(P_ε) : Find $p_\varepsilon \in V$ such that

$$\langle A(p_\varepsilon), q \rangle + \frac{1}{\varepsilon} \langle p_\varepsilon^-, q \rangle = \langle f, q \rangle, \quad \forall q \in V \quad (3.27)$$

where ε is positive.

The mechanism implied in (3.27) is that if the solution p_ε can approach the constraint set K when $\varepsilon \rightarrow 0$, and if the second term in (3.27) has reduced dramatically, then (3.27) approximates (3.3) with p_ε approximating p .

3.4 TREATMENT OF THE FREE BOUNDARY.

In numerical analysis, lubrication engineers traditionally employ an iterative routine to locate the free boundary, where the boundary is moved each iteration. The idea is to simply set the negative part of p to be zero, or to adjust the domain of integration of p when p appears negative. It is not clear that this adjustment process is always convergent.

However, the penalty method leads to a regularized formulation that does admit to some mathematical analysis. We shall prove the convergence of penalty method in Chapter 4 and show some more results about approximations in finite-dimensional spaces in Chapter 6.

4. QUALITATIVE ANALYSIS

While elastohydrodynamic lubrication theory has been available to tribologists for decades, the mathematical theory of the governing equations and inequalities is in its very early stages of development. Capriz and Cimatti^[11] contributed a significant survey of this theory and described some preliminary results on the existence, uniqueness and regularity of solutions (see [10,13,15,16,17]). Although these results are principally for the Reynolds equation without elastic deformation, Cimatti^[18] did generalize some of the ideas to a class of nonlinear problems. By using the theory of nonlinear variational inequalities with pseudomonotone operators, Oden and Wu^[62] succeeded in proving the existence of solutions for the general nonlinear Reynolds–Hertz equations. Wu^[93] further advanced the study by both proving the convergence of a penalty method and establishing the regularity of the solutions for the line contact case.

In this chapter, we use the results and definitions of the preceding chapters to study properties of the operator A and to prove the existence of weak solutions to Reynold–Hertz equations for elastohydrodynamic lubrication. These represent extensions of previous work^[62]. In addition, we shall prove the existence of solutions of penalized problems and the convergence of penalty method.

Another important result established in this chapter is the regularity of the line contact solutions. As a consequence of this regularity, we are able to prove the validity of Reynolds condition (2.21) and thus we are able to prove the existence of solutions in classical sense.

4.1 PSEUDOMONOTONE OPERATORS.

By Theorem 3.1 to 3.3, the pseudomonotonicity of an operator is important to the theory of variational inequalities. However, to prove pseudomonotonicity it is useful to establish some stronger sufficient conditions, cited from [55].

THEOREM 4.1. Let (3.17) hold. Let K be unbounded and operator A

be bounded and coercive in the sense of (3.12). Operator A is pseudomonotone if one set of the following conditions is satisfied:

(1) U is reflexive Banach space and A is hemicontinuous and monotone on K .

(2) A is hemicontinuous and there exists a completely continuous operator

$B : U \rightarrow U'$ such that

$$\langle A(u) - A(v), u - v \rangle \geq \langle B(u) - B(v), u - v \rangle, \quad \forall u, v \in K \quad (4.1)$$

Complete continuity means that if $u_n \rightarrow u$ weakly in U , then $B(u_n) \rightarrow B(u)$ strongly in U' .

(3) A is a Gårding-type operator, i.e., A is hemicontinuous, and for a Banach space W in which U is compactly embedded there exists a continuous, nonnegative function $H : R^+ \times R^+ \rightarrow R^+$ with the property

$$\lim_{\theta \rightarrow 0^+} \frac{1}{\theta} H(x, \theta y) = 0, \quad \forall x, y \in R^+ \quad (4.2)$$

and A satisfies

$$\langle A(u) - A(v), u - v \rangle \geq -H(\mu, \|u - v\|_W), \quad \forall u, v \in B_\mu(0) \cap K \quad (4.3)$$

where $B_\mu(0) = \{w \in U \mid \|w\|_U < \mu, \mu > 0\}$ is the open ball in U

4.2 PROPERTIES OF OPERATOR A .

To fix our idea, we begin by considering only a classical Newtonian fluid (the case $\alpha=0$). In other words, we take $\mu=\mu_0$ and the factor $e^{-\alpha p}$ drops from the equation. For the general case, we may transform p to an reduced pressure

$\tilde{p} = (1 - e^{-\alpha p})/\alpha$ and $\nabla \tilde{p} = e^{-\alpha p} \nabla p$. We shall return to this case at the end of this section.

LEMMA 4.2. For $h_1(p)$ defined in (2.29) and (2.30), $0 < \delta < 1$, and

$q=(2-\delta)/(1-\delta) > 2$, there exists a constant $C_h > 0$ such that

$$\sup_{(x,y) \in \Omega} |h_1(p)| \leq C_h \|p\|_{L^q} \quad \forall p \in V$$

$$h_1(p)(x,y) \in L^\infty(\Omega) \quad (4.4)$$

PROOF : Rewrite (2.29) and (2.30) in the form of a kernel integral

$$h_1(p) = C_E \int_{\Omega} K(x,\xi) p(\xi) d\Omega, \quad C_E = 2/\pi E' \quad (4.5)$$

For $x \in \Omega$, in case (1) (line contact),

$$K = \ln(x_0 - \xi / x - \xi)^2 \in L^s(\Omega), \quad s > 0 \quad (4.6)$$

and in case (2) (point contact)

$$K = r^{-1} \in L^{2-\delta}(\Omega), \quad 0 < \delta < 1, \quad r = \sqrt{(x-\xi)^2 + (y-\eta)^2} \quad (4.7)$$

By Hölder's inequality, we have always

$$|h_1(p)| \leq C_E \|K\|_{L^{2-\delta}(\Omega)}(x,y) \|p\|_{L^q}, \quad 0 < \delta < 1 \quad (4.8)$$

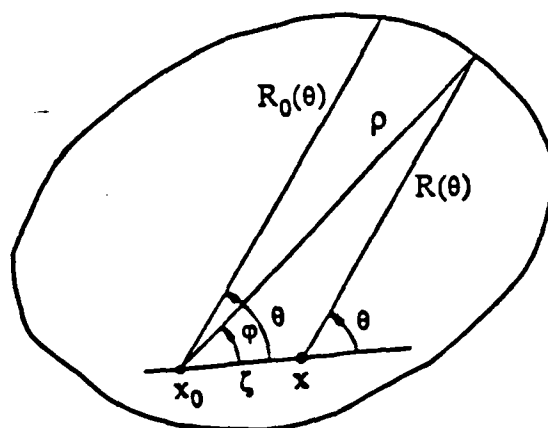
It is evident that in case (1) $\|K\|_{L^{2-\delta}}$ is a continuous function of (x,y) in $\bar{\Omega}$. In case (2), denote

$$\begin{aligned} I(x,y) &= (\|K\|_{L^{2-\delta}})^{2-\delta} = \int_{\Omega(\xi,\eta)} |K(x,y,\xi,\eta)|^{2-\delta} d\xi d\eta \\ &= \int_0^{2\pi} d\theta \int_0^{R(\theta)} \frac{r dr}{r^{2-\delta}} = \frac{1}{\delta} \int_0^{2\pi} R^\delta(x,y,\theta) d\theta, \quad \text{case (a): } (x,y) \in \text{int } \Omega \\ &= \int_0^\pi d\theta \int_0^{R(\theta)} \frac{r dr}{r^{2-\delta}} = \frac{1}{\delta} \int_0^\pi R^\delta(x,y,\theta) d\theta, \quad \text{case (b): } (x,y) \in \partial\Omega \end{aligned} \quad (4.9)$$

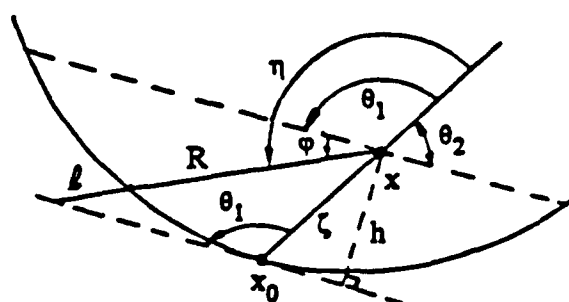
It is shown below that the function $I(x,y)$ in (4.9) is also continuous on $\bar{\Omega}$. Then if we denote $M_K = \max \|K\|_{L^{2-\delta}}$, we have (4.4).

With the notations shown in Fig. 4.1,

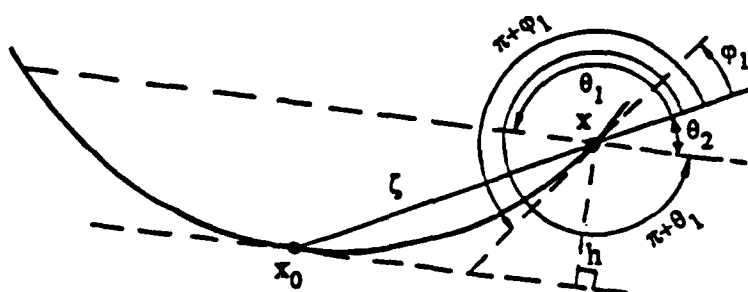
$$(a) \quad (x_0, y_0) \in \text{int } \Omega. \quad \text{Denote } J(x_0, y_0) = \int_0^{2\pi} (R_0)^\delta d\theta \quad \text{and } J(x,y) = \int_0^{2\pi} R^\delta d\theta$$



(a) Case (a) : $x, x_0 \in \text{int } \Omega$



(b) Case (b₁) : $x \in \text{int } \Omega, x_0 \in \partial \Omega$



(c) Case (b₂) : $x \in \partial \Omega, x_0 \in \partial \Omega$

Fig. 4.1 Analysis of the Distance x_0x

with θ measured in a same manner from the line x_0x . $R^2 = \rho^2 + \zeta^2 - 2\rho\zeta \cos \phi$.

For (x,y) close to (x_0,y_0) , $\zeta = o(\rho)$, $|\theta - \phi| = O(\zeta)$. So $\rho = R_0(\phi) = R_0(\theta) + O(\zeta)$,

$R^2 = R_0^2 + O(\zeta)$, $R^\delta = R_0^\delta + O(\zeta^\delta)$. Thus $J(x,y) = J(x_0,y_0) + O(\zeta^\delta)$, so in case

(a), the function in (4.9) is continuous at $(x,y) \in \text{int } \Omega$.

(b) $(x_0,y_0) \in \partial\Omega$. Consider a neighboring point (x,y) .

(b₁) If $(x,y) \in \text{int } \Omega$, see Fig.4.1b, we still measure θ from the line x_0x .

Then

$$J(x,y) = \int_0^{2\pi} R^\delta(x,y,\theta) d\theta, \quad J(x_0,y_0) = \left(\int_0^{\theta_1} + \int_{2\pi-\theta_2}^{2\pi} \right) R^\delta(x,y,\theta) d\theta.$$

$$J(x,y) - J(x_0,y_0) = \left(\int_0^{\theta_1} + \int_{2\pi-\theta_2}^{2\pi} \right) (R^\delta - R_0^\delta) d\theta + \int_{\theta_1}^{2\pi-\theta_2} R^\delta d\theta.$$

The same argument as in part (a) leads to

$$\left(\int_0^{\theta_1} + \int_{2\pi-\theta_2}^{2\pi} \right) (R^\delta - R_0^\delta) d\theta = O(\zeta^\delta).$$

For $\theta_1 < \eta < 2\pi - \theta_2 = \pi + \theta_1$, $R \leq h/\sin \psi$, $\psi = \eta - \theta_1 \in [0, \pi]$, $h \leq \zeta$.

$$\begin{aligned} \left| \int_{\theta_1}^{2\pi-\theta_2} R^\delta d\theta \right| &\leq h^\delta \int_0^\pi \frac{d\psi}{|\sin \psi|^\delta} = 2h^\delta \int_0^{\pi/2} \frac{d\psi}{\sin^\delta \psi} \\ &\leq 2h^\delta \int_0^{\pi/2} \frac{d\psi}{\psi^\delta (2/\pi)^\delta} = O(\zeta^\delta) \end{aligned}$$

Therefore $|J(x,y) - J(x_0,y_0)| = O(\zeta^\delta)$.

(b₂) If $(x,y) \in \partial\Omega$, see Fig.4.1c. Suppose $\theta_2 < \pi/2$, $\phi_1 < \theta_1 < \pi + \phi_1$.

Let $\Delta = J(x,y) - J(x_0,y_0) = \Delta_1 + \Delta_2 + \Delta_3$, $\Delta_1 = \int_{\phi_1}^{\theta_1} (R^\delta - R_0^\delta) d\theta$, $\Delta_2 = \int_{\theta_1}^{\pi+\phi_1} R^\delta d\theta$,

$\Delta_3 = \int_{\pi+\theta_1}^{2\pi} R_0^\delta d\theta$. Similar to part (a) and (b₁), $\Delta_1 = O(\zeta^\delta)$, $\Delta_2 = O(\zeta^\delta)$.

Then for $\pi + \theta_1 < \theta < 2\pi$, $\eta = \theta - \pi - \theta_1 \in (0, \theta_2)$, $R < d/\cos\eta < d/\cos\theta_2 = \zeta \Rightarrow \Delta_3 = O(\zeta^{\delta/2})$.

Hence $\Delta = O(\zeta^{\delta/2})$.

In summary we have proved that $I(x, y) - I(x_0, y_0) = O(\zeta^{\delta/2})$ and therefore $\|K\|_{L^{2-\delta}}$ is continuous. #

COROLLARY 4.3. $h_1(p)$ and $h(p)$ are bounded functions if $p \in L^q(\Omega)$, and in fact,

$$\tilde{h}_1(p) \in L^\infty(\Omega), \quad h(p) \in L^\infty(\Omega) \quad (4.10)$$

PROOF : In case (2.28), it is obvious. In case (2.28') we have $s(h) = \inf(h_2 + h_1) \leq (h_2 + h_1)|_{(x,y)=(0,0)} = h_1|_{(0,0)} \Rightarrow |s(h)| \leq \sup |h_1(p)|$. We still have $\tilde{h}_1(p) \in L^\infty(\Omega)$ and $h \in L^\infty(\Omega)$. Also $\sup |\tilde{h}_1| < C_h \|p\|_{L^q}$, with $C_h = 2M_k$. #

THEOREM 4.4. Operator A is bounded.

PROOF : By virtue of the duality, for any $q \in V$,

$$\begin{aligned} \langle A(p), q \rangle_{V_X V} &= \langle -\nabla \cdot (h^3(p) \nabla p) + 12\mu_0 \partial(u_0 \tilde{h}_1(p)) / \partial x, q \rangle \\ &= \langle h^3(p) \nabla p, \nabla q \rangle - \langle 12\mu_0 u_0 \tilde{h}_1(p), \partial q / \partial x \rangle \\ &= \int_{\Omega} (h^3(p) \nabla p \cdot \nabla q - 12\mu_0 u_0 \tilde{h}_1(p) \partial q / \partial x) d\Omega \end{aligned}$$

$h(p) = \tilde{h}_1 + \tilde{h}_2$, u_0 and h_2 are smooth. Denote

$$M_u = \sup_{\Omega} |u_0|, \quad M_h = \sup_{\Omega} |\tilde{h}_2| \quad (4.11)$$

Then by Hölder's inequality

$$\begin{aligned} \int \left| \frac{\partial q}{\partial x} \right| d\Omega &\leq \int |\nabla q| d\Omega \leq \left(\int |\nabla q|^2 d\Omega \right)^{1/2} \left(\int 1 d\Omega \right)^{1/2} \\ &\leq C_{\Omega} \|q\|_{H^1}, \quad C_{\Omega} = (\text{meas } \Omega)^{1/2} \\ |\langle A(p), q \rangle| &\leq \sup_{\Omega} |h^3(p)| \left(\int_{\Omega} |\nabla p|^2 d\Omega \right)^{1/2} \left(\int_{\Omega} |\nabla q|^2 d\Omega \right)^{1/2} \\ &\quad + 12\mu_0 M_u \sup |\tilde{h}_1(p)| \int_{\Omega} |\partial q / \partial x| d\Omega \end{aligned} \quad (4.12)$$

$$\leq (M_h + C_h \|p\|_{L^q})^3 \|p\|_{H^1} \|q\|_{H^1} + 12\mu_0 M_u C_h \|p\|_{L^q} C_\Omega \|q\|_{H^1}$$

By the Sobolev embedding theorem (see [1], for $n=2$, $m=1$, $p=2$ case, $mp=n$), we have $H^1(\Omega) \subset L^q(\Omega)$ compactly. Thus there exists a constant $C_q > 0$ such that

$$\|p\|_{L^q} \leq C_q \|p\|_{H^1}, \quad \forall p \in H^1(\Omega) \quad (4.13)$$

Then $\|A(p)\|_V \leq (M_h + C_h C_q \|p\|_{H^1})^3 + 12\mu_0 M_u C_h C_q C_\Omega \|p\|_{H^1}$. This means that A maps bounded sets in V into bounded sets in V , so A is bounded. #

THEOREM 4.5. Operator A is coercive on V in the sense of (3.12) with the choice of $p_0=0$. In fact, there exist $C_1, C_2 > 0$ such that

$$\langle A(p), p \rangle \geq C_1 \|p\|_1^2 - C_2 \|p\|_1, \quad \forall p \in V \quad (4.14)$$

PROOF : From (2.28'), we have $\tilde{h}_2 = h_z + h_m \geq h_m > 0$ and $h \geq h_m$.

$$\langle A(p), p \rangle = \int_\Omega (h^3(p) |\nabla p|^2 - 12\mu_0 u_0 \tilde{h}_1(p) \partial p / \partial x) d\Omega.$$

Let

$$\Omega_1 = \{(x,y) \in \Omega \mid \tilde{h}_1(p) \geq 0\}; \quad \Omega_2 = \{(x,y) \in \Omega \mid \tilde{h}_1(p) < 0\}.$$

Note that $|\partial p / \partial x| \leq |\nabla p|$. Let

$$\Omega_{11} = \{(x,y) \in \Omega_1 \mid 12\mu_0 u_0 |\nabla p| < h^2 |\nabla p|^2\}; \quad \Omega_{12} = \Omega_1 - \Omega_{11}$$

$$\begin{aligned} \int_{\Omega_{11}} (\dots) d\Omega &= \int_{\Omega_{11}} [\tilde{h}_2 h^2(p) |\nabla p|^2 + \tilde{h}_1(p) (h^2(p) |\nabla p| - 12\mu_0 u_0 \partial p / \partial x)] d\Omega \\ &\geq \int_{\Omega_{11}} \tilde{h}_2 h^2(p) |\nabla p|^2 d\Omega \end{aligned}$$

In Ω_{12} , $h^2 |\nabla p|^2 \leq 12\mu_0 u_0 |\nabla p|$, $|\nabla p| \leq 12\mu_0 u_0 h^{-2} \leq 12\mu_0 M_u h_m^{-2}$.

$$\begin{aligned} \int_{\Omega_{12}} (\dots) d\Omega &\geq \int_{\Omega_{12}} [h^3(p) |\nabla p|^2 - 12\mu_0 u_0 \tilde{h}_1(p) |\nabla p|] d\Omega \\ &\geq \int_{\Omega_{12}} [h^3 |\nabla p|^2 - 12\mu_0 u_0 h_1(p) \cdot 12\mu_0 M_u h_m^{-2}] d\Omega \\ &\geq \int_{\Omega_{12}} h^3 |\nabla p|^2 d\Omega - (12\mu_0 M_u h_m^{-1})^2 C_h \|p\|_{L^q} \text{meas}(\Omega_{12}). \end{aligned}$$

In Ω_2 , $\tilde{h}_1 < 0$. Let

$$\Omega_{22} = \{(x,y) \in \Omega_2 \mid \partial p / \partial x \geq 0\}, \quad \Omega_{21} = \Omega_2 - \Omega_{22}.$$

In Ω_{22} , $\tilde{h}_1 \partial p / \partial x < 0$, and then

$$\int_{\Omega_{22}} (\dots) d\Omega \geq \int_{\Omega_{22}} h^3(p) |\nabla p|^2 d\Omega$$

In Ω_{21} , $\partial p / \partial x \geq -|\nabla p|$, $\tilde{h}_1 < 0$, we have

$$\int_{\Omega_{21}} (\dots) d\Omega \geq \int_{\Omega_{21}} [h^3(p) |\nabla p|^2 + 12\mu_0 u_0 (h(p) - \tilde{h}_2) |\nabla p|] d\Omega$$

$h(p) > 0$ and $|\tilde{h}_2| \leq M_h$, so we have

$$\int_{\Omega_{21}} (\dots) d\Omega \geq \int_{\Omega_{21}} h^3(p) |\nabla p|^2 - 12\mu_0 M_u M_h \int_{\Omega_{21}} |\nabla p| d\Omega.$$

In summary, we obtain

$$\begin{aligned} \langle A(p), p \rangle &\geq \int_{\Omega_{11}} \tilde{h}_2 h^2(p) |\nabla p|^2 d\Omega + \int_{\Omega_{12}} h^3(p) |\nabla p|^2 d\Omega \\ &\quad - (12\mu_0 M_u h_m^{-1})^2 C_h C_q \text{meas}(\Omega_{12}) \|p\|_{H^1} \\ &\quad + (\int_{\Omega_{21}} + \int_{\Omega_{22}}) h^3(p) |\nabla p|^2 d\Omega - 12\mu_0 M_u M_h C_{\Omega_{21}} \|p\|_{H^1} \\ &\geq h_m^3 (\|p\|_{H^1})^2 - C_2 \|p\|_{H^1}. \end{aligned}$$

Here $\|\cdot\|_{H^1}$ is the seminorm.

By Poincaré's inequality, there exists $C_0 > 0$, such that

$$(\|p\|_{H^1})^2 \leq C_0 (\|p\|_{H^1})^2 \quad \forall p \in H^1_0(\Omega) \quad (4.15)$$

Thus we have $\langle A(p), p \rangle \geq C_1 (\|p\|_{H^1})^2 - C_2 \|p\|_{H^1}$.

Here we define

$$C_1 = h_m^3 / C_0$$

$$C_2 = (12\mu_0 M_u h_m^{-1})^2 C_h C_q \text{meas}(\Omega) + 12\mu_0 M_u M_h C_{\Omega} \quad (4.16)$$

Therefore

$$\lim_{\|p\|_{H^1} \rightarrow \infty} \frac{\langle A(p), p \rangle}{\|p\|_{H^1}} = +\infty \quad \#$$

Note, in [62] we used the expression (2.28) and supposed $h_0 > 0$. In fact, in

the case of heavy loads, h_0 might be negative. Now we have extended the analysis to a general case with the expression (2.28').

THEOREM 4.6. Operator A is hemicontinuous. That is

$$\lim_{t \rightarrow 0+} \langle A(p+ tq), r \rangle = \langle A(p), r \rangle, \quad \forall p, q, r \in V \quad (4.17)$$

PROOF: $\langle A(p+ tq), r \rangle$

$$= \int_{\Omega} [(h^3(p+q) \nabla(p+q) \cdot \nabla r - 12\mu_0 u_0 \tilde{h}_1(p+q) \partial r / \partial x)] d\Omega$$

$h(p) = \tilde{h}_2 + \tilde{h}_1(p) = h_z + h_m + h_1(p) - s(h(p))$. $\tilde{h}_1(p)$ is linear in p , $s(h) = \min(h_z, h_1)$ is continuous in h_1 , and continuous in p . The other terms are all continuous in the parameter t . So we have

$$\lim_{t \rightarrow 0+} \langle A(p+ tq), r \rangle = \int_{\Omega} [(h^3(p) \nabla p \cdot \nabla r - 12\mu_0 u_0 \tilde{h}_1(p) \partial r / \partial x)] d\Omega = \langle A(p), r \rangle. \quad \#$$

THEOREM 4.7. Denote $B_{\eta} = \{q \in V \mid \|q\|_{H^1} \leq \eta\}$, a ball in V . There

exists a constant $C_{\eta}(\eta, M_u, C_h, \Omega) > 0$ such that $\forall p, q \in B_{\eta}$,

$$\langle A(p) - A(q), p - q \rangle \geq C_1 \|p - q\|_1^2 - C_{\eta} \|p - q\|_{L^q} \|p - q\|_{H^1} \quad (4.18)$$

with C_1 , the constant appearing in (4.14) and defined in (4.16). For simplicity, we denote $\|\cdot\|_1$ for $\|\cdot\|_{H^1}$, and $|\cdot|_1$ for $|\cdot|_{H^1}$.

PROOF: $\langle A(p) - A(q), p - q \rangle$

$$\begin{aligned} &= \int_{\Omega} [h^3(p) \nabla p \cdot \nabla(p - q) - h^3(q) \nabla q \cdot \nabla(p - q) \\ &\quad - 12\mu_0 u_0 (\tilde{h}_1(p) - \tilde{h}_1(q)) \partial(p - q) / \partial x] d\Omega \\ &= \int_{\Omega} [h^3(p) \nabla(p - q) \cdot \nabla(p - q) + (h^3(p) - h^3(q)) \nabla q \cdot \nabla(p - q) \\ &\quad - 12\mu_0 u_0 (\tilde{h}_1(p) - \tilde{h}_1(q)) \partial(p - q) / \partial x] d\Omega \end{aligned} \quad (4.19)$$

The first integral,

$$\int_{\Omega} h^3(p) |\nabla(p - q)|^2 d\Omega \geq h_m^3 |p - q|_1^2 \geq C_1 \|p - q\|_1^2.$$

The constant C_1 is defined in (4.16).

Note that $\tilde{h}_1(p) - \tilde{h}_1(q) = h_1(p) - h_1(q) - (s(h(p)) - s(h(q)))$. Since $s(h(p)) = \min(h_2 + h_1(p))$, $|s(h(p)) - s(h(q))| \leq \sup |(h_2 + h_1(p)) - (h_2 + h_1(q))|$, $|h_1(p) - h_1(q)| = |\tilde{h}_1(p) - \tilde{h}_1(q)| \leq C_h \|p - q\|_{L^q}$. Thus

$$|\tilde{h}_1(p) - \tilde{h}_1(q)| \leq 2 \sup |h_1(p) - h_1(q)| \leq 2C_h \|p - q\|_{L^q} \quad (4.20)$$

So for the third integral in (4.19), we have

$$\begin{aligned} & \left| \int_{\Omega} -12\mu_0 u_0 (\tilde{h}_1(p) - \tilde{h}_1(q)) \partial(p - q) / \partial x \, d\Omega \right| \\ & \leq 24\mu_0 M_u C_h \|p - q\|_{L^q} \int_{\Omega} |\nabla(p - q)| \, d\Omega \\ & \leq 24\mu_0 M_u C_h C_{\Omega} \|p - q\|_{L^q} \|p - q\|_{H^1} \end{aligned}$$

Here M_u , C_h and C_{Ω} are defined in (4.11), (4.4) and (4.12), respectively.

For the second integral in (4.19),

$$\begin{aligned} h^3(p) - h^3(q) &= (h(p) - h(q)) (h^2(p) + h(p)h(q) + h^2(q)) \\ |h(p)| &\leq M_h + C_h \|p\|_{L^q} \leq M_h + C_h C_q \|p\|_{H^1} \leq M_h + C_h C_q \eta, \end{aligned}$$

also $|h(q)| \leq M_h + C_h C_q \eta$. Where C_q is defined in (4.13). Recalling (4.20), $|h(p) - h(q)| \leq 2C_h \|p - q\|_{L^q}$, we have

$$\begin{aligned} & \left| \int_{\Omega} (h^3(p) - h^3(q)) \nabla q \cdot \nabla(p - q) \, d\Omega \right| \\ & \leq 2C_h \|p - q\|_{L^q} 3(M_h + C_h C_q \eta)^2 \int_{\Omega} |\nabla q| \cdot |\nabla(p - q)| \, d\Omega \\ & \leq C_3 \|p - q\|_{L^q} \|q\|_{H^1} \|p - q\|_{H^1} \leq C_3 \eta \|p - q\|_{L^q} \|p - q\|_{H^1} \end{aligned}$$

Here $C_3 = 6C_h (M_h + C_h C_q \eta)^2$. In summary, we obtain

$$\langle A(p) - A(q), p - q \rangle \geq C_1 \|p - q\|_1^2 - (24\mu_0 M_u C_h C_{\Omega} + C_3 \eta) \|p - q\|_{L^q} \|p - q\|_{H^1}$$

We thus need only denote

$$C_{\eta} = 24\mu_0 M_u C_h C_{\Omega} + 6C_h (M_h + C_h C_q \eta)^2 \eta \quad (4.21)$$

to obtain the desired result. #

Inequality (4.18) is similar to the inequality (4.3) appearing in the definition of a Gårding operator except that we now have a function H in the form of

$H(\mu, \|u-v\|_w \|u-v\|_v)$. Thus we need only retrace the steps used in proving a Gårding operator to be pseudomonotone, although the procedure is not without some difficulties. In fact we may consider operator A to be a kind of generalized Gårding operator.

THEOREM 4.8. Operator A is pseudomonotone.

PROOF: By Definition 3.8, we need to examine the behavior of A with a sequence $\{p_n\}$ satisfying property P (3.15). Now suppose $\{p_n\} \in V$ and $p_n \rightarrow p$ weakly in V , and

$$\overline{\lim} \langle A(p_n), p_n - p \rangle \leq 0 \quad (4.22)$$

we are to prove $\forall q \in V = H^1_0(\Omega)$, $\underline{\lim} \langle A(p_n), p_n - q \rangle \geq \langle A(p), p - q \rangle$.

(1) Since $p_n \rightarrow p$ weakly in $V = H^1_0(\Omega)$, the reflexive Banach space, $\{p_n\}$ must be bounded in V . Let $\|p_n\|_{H^1} \leq \eta$, $\|p\|_{H^1} \leq \eta$, or we say $\{p_n\}, p \in B_\eta$.

According to the Sobolev embedding theorem, H^1 is embedded in L^q compactly, so $p_n \rightarrow p$ strongly in L^q . Thus we have $\langle A(p), p_n - p \rangle \rightarrow 0$, and $\|p_n - p\|_{L^q} \rightarrow 0$ as $n \rightarrow \infty$. Hence by the boundedness of $\{p_n\}$ and from (4.18) and (4.22), we obtain

$$\begin{aligned} 0 &\geq \overline{\lim} \langle A(p_n), p_n - p \rangle \geq \underline{\lim} \langle A(p_n), p_n - p \rangle \\ &\geq \underline{\lim} [\langle A(p), p_n - p \rangle - C_\eta \|p_n - p\|_{L^q} \|p_n - p\|_{H^1}] = 0. \end{aligned}$$

Therefore

$$\lim \langle A(p_n), p_n - p \rangle = 0 \quad (4.23)$$

(2) $\forall r \in V$, we use essentially the process used in proving Lemma 4.7 to obtain

$$\begin{aligned} \langle A(p_n) - A(r), p_n - r \rangle &= \int_\Omega [h^3(p_n) \nabla(p_n - r) \cdot \nabla(p_n - r) \\ &\quad + (h^3(p_n) - h^3(r)) \nabla r \cdot \nabla(p_n - r) - 12\mu_0 u_0 (\tilde{h}_1(p_n) - \tilde{h}_1(r)) \partial(p_n - r) / \partial x] d\Omega \end{aligned}$$

$$\begin{aligned} &\geq \int_{\Omega} [(h(p_n) - h(p) + h(p))^3 - h^3(r)] \nabla r \cdot \nabla (p_n - r) \\ &\quad - 12\mu_0 u_0 (\tilde{h}_1(p_n) - \tilde{h}_1(r)) \partial(p_n - r) / \partial x] d\Omega \end{aligned} \quad (4.24)$$

We next split the integral in (4.24) into three parts, $I_1 + I_2 + I_3$, where

$$\begin{aligned} I_1 = &\int_{\Omega} [(h(p_n) - h(p)) \cdot ((h(p_n) - h(p))^2 + 3(h(p_n) - h(p)) h(p) + 3h^2(p)) \nabla r \cdot \nabla (p_n - r) \\ &- 12\mu_0 u_0 (\tilde{h}_1(p_n) - \tilde{h}_1(p)) \partial(p_n - r) / \partial x] d\Omega \end{aligned}$$

$$I_2 = \int_{\Omega} [(h^3(p) - h^3(r)) \nabla r \cdot \nabla (p_n - p) - 12\mu_0 u_0 (\tilde{h}_1(p) - \tilde{h}_1(r)) \partial(p_n - p) / \partial x] d\Omega$$

$$I_3 = \int_{\Omega} [(h^3(p) - h^3(r)) \nabla r \cdot \nabla (p - r) - 12\mu_0 u_0 (\tilde{h}_1(p) - \tilde{h}_1(r)) \partial(p - r) / \partial x] d\Omega$$

Using (4.20), we have

$$|h(p_n) - h(p)| = |\tilde{h}_1(p_n) - \tilde{h}_1(p)| \leq 2C_h \|p_n - p\|_{L^q} \rightarrow 0$$

Because of the fact that $\{p_n\}$ is bounded, $\{h(p_n)\}$ and $\{\tilde{h}_1(p_n)\}$ are bounded.

Then $I_1 \rightarrow 0$ when $n \rightarrow \infty$. Moreover,

$$\begin{aligned} I_2 = &\langle -\nabla \cdot ((h^3(p) - h^3(r)) \nabla r), p_n - p \rangle \\ &+ \langle 12\mu_0 \partial(u_0(\tilde{h}_1(p) - \tilde{h}_1(r))) / \partial x, p_n - p \rangle \rightarrow 0, \end{aligned}$$

because of the weak convergence of $p_n \rightarrow p$. Finally, I_3 is the sum of the second and third integrals in (4.19). Thus, following the same reasoning as that leading to (4.19) and passing to the limit, we have

$$\lim \langle A(p_n) - A(r), p_n - r \rangle \geq I_3 \geq -C_{\eta} \|p - r\|_{L^q} \|p - r\|_{H^1} \quad (4.25)$$

(3) In view of (4.23) and (4.25)

$$\begin{aligned} \lim \langle A(p_n), p - r \rangle &= \lim \langle A(p_n), p_n - r \rangle \\ &= \lim \langle A(p_n) - A(r) + A(r), p_n - r \rangle \\ &\geq \lim \langle A(r), p_n - r \rangle - C_{\eta} \|p - r\|_{L^q} \|p - r\|_{H^1} \\ &= \langle A(r), p - r \rangle - C_{\eta} \|p - r\|_{L^q} \|p - r\|_{H^1} \end{aligned} \quad (4.26)$$

(4) Now for any $q \in V$, set $r = p + \theta(q - p) \in V$, $0 < \theta < 1$, for small θ , $\|r\| < \eta$,

since $p \in B_\eta$. Then $p-r = \theta(p-q)$. Substitution for r into (4.26) gives the result

$$\lim_{\theta \rightarrow 0+} \langle A(p_\eta), \theta(p-q) \rangle \geq \langle A(r), \theta(p-q) \rangle - C_\eta \theta^2 \|p-q\|_{L^q} \|p-q\|_{H^1}$$

Dividing by θ and using the hemicontinuity of A , we get

$$\lim_{\theta \rightarrow 0+} \langle A(p+\theta(p-q)), p-q \rangle = \langle A(p), p-q \rangle$$

We have finally

$$\lim_{\theta \rightarrow 0+} \langle A(p_\eta), p-q \rangle \geq \langle A(p), p-q \rangle \quad \forall q \in V$$

This is precisely the requirement that A be a pseudomonotone operator. #

THEOREM 4.9. Operator A is continuous, in fact, uniformly continuous in ball B_η , i.e., there exists a constant $C_A(\eta) > 0$ such that

$$\|A(p) - A(q)\|_{H^{-1}} \leq C_A \|p-q\|_{H^1}, \quad \forall p, q \in V, \quad \|p\|_v, \|q\|_v < \eta \quad (4.27)$$

PROOF: Referring to (4.19), a similar analysis for $\langle A(p) - A(q), r \rangle$ leads to the inequality

$$\begin{aligned} |\langle A(p) - A(q), r \rangle| &\leq K_h^3 \|p-q\|_{H^1} \|r\|_{H^1} + 3K_h^2 2C_h \|p-q\|_{L^q} \|q\|_{H^1} \|r\|_{H^1} \\ &\quad + 12\mu_0 M_u 2C_h C_\Omega \|p-q\|_{L^q} \|r\|_{H^1} \end{aligned}$$

Here we denote $K_h = |\sup(h(p))| \leq M_h + C_h \|p\|_{L^q} \leq M_h + C_h C_q \eta$, and then

$$C_A = K_h^3 + 6K_h^2 C_h C_q \eta + 24\mu_0 M_u C_h C_\Omega \quad (4.28)$$

At last, $\|A(p) - A(q)\|_{H^{-1}} \leq C_A \|p-q\|_{H^1}$. #

REMARK 1. Physically p must be bounded (e.g., by the yield stress) and $e^{-\alpha p}$ must be bounded below from zero. However an assumption of boundedness for p is not appropriate or necessary to our analysis. We prefer to modify the viscosiry as : $\mu = \mu_0 \mu_1$, μ_0 - classical viscosity

$$\mu_1 = \begin{cases} e^{\alpha p_M} & \text{if } p \geq p_M > 0 \\ e^{\alpha p} & \text{if } -p_M < p < p_M \\ e^{-\alpha p_M} & \text{if } p \leq -p_M \end{cases} \quad (4.29)$$

Then define \tilde{p} so that $\nabla \tilde{p} = \nabla p / \mu_1$ appears in Reynolds equation (2.23)

$$\tilde{p} = \begin{cases} e^{-\alpha p_M} (p - p_M) + (1 - e^{-\alpha p_M}) / \alpha & \text{if } p \geq p_M \\ (1 - e^{-\alpha p_M}) / \alpha & \text{if } |p| < p_M \\ e^{\alpha p_M} (p + p_M) + (1 - e^{-\alpha p_M}) / \alpha & \text{if } p \leq -p_M \end{cases} \quad (4.30)$$

$$p = \begin{cases} p_M + e^{\alpha p_M} (\tilde{p} - (1 - e^{-\alpha p_M}) / \alpha) & \text{if } \tilde{p} \geq (1 - e^{-\alpha p_M}) / \alpha (>0) \\ -\ln (1 - e^{-\alpha p}) / \alpha & \text{if } (1 - e^{-\alpha p_M}) / \alpha < \tilde{p} < (1 - e^{-\alpha p_M}) / \alpha \\ -p_M + e^{-\alpha p_M} (\tilde{p} - (1 - e^{-\alpha p_M}) / \alpha) & \text{if } \tilde{p} \leq (1 - e^{-\alpha p_M}) / \alpha (<0) \end{cases} \quad (4.31)$$

It is easy to verify that $|p| < C_p |\tilde{p}|$ with $C_p = \max(p_M, e^{\alpha p_M})$; $\|p\|_{L^q} \leq C_p \|\tilde{p}\|_{L^q}$.

Then we still have some constant $C_h = C_h C_p$ so that $|h_1(\tilde{p})| = |h_1(p(\tilde{p}))| \leq C_h \|\tilde{p}\|_{L^q}$. By the continuity of $p(\tilde{p})$, there is a $C_p > 0$ such that $|p_1 - p_2| = |p(\tilde{p}_1) - p(\tilde{p}_2)| \leq C_p \|\tilde{p}_1 - \tilde{p}_2\|$; and $|\tilde{h}_1(\tilde{p}_1) - \tilde{h}_1(\tilde{p}_2)| \leq 2C_h \|p(\tilde{p}_1) - p(\tilde{p}_2)\|_{L^q} \leq 2C_h C_p \|\tilde{p}_1 - \tilde{p}_2\|_{L^q}$

For simplicity we keep the notations used before. In later sections, we proceed with $\alpha=0$, unless indicated otherwise.

REMARK 2. All the analyses here are valid for any closed convex subset of $V = H^1_0(\Omega)$.

4.3 EXISTENCE THEOREM.

In Section 4.2 we proved that the operator A is bounded, coercive and pseudomonotone. According to Theorem 3.3, (3.24) has at least one solution. Furthermore by Theorem 3.6, this solution is also a weak solution of (3.3). We state this conclusion as

THEOREM 4.10. For the nonlinear free boundary problem in elastohydrodynamic lubrication, there exists at least one solution to the variational inequality (3.24). Any such solution is a weak solution of the Reynolds-Hertz equations (2.23-25). #

4.4 BRIEF COMMENTS ON UNIQUENESS.

As indicated in Corollary 3.5, if in addition to the conditions listed in Theorem 4.10, A is strictly monotone, then the solution of (3.24) is unique.

We note that (4.18) is of the form of a perturbation of a monotone operator. Using (4.13) : $\|p-q\|_{L^q} \leq C_q \|p-q\|_{H^1}$, we go a step further to obtain

$$\langle A(p) - A(q), p - q \rangle \geq (C_1 - C_\eta C_q) (\|p - q\|_{H^1})^2 \quad (4.32)$$

Thus if we have in addition, the following condition

$$C_1 - C_\eta C_q > 0 \quad (4.33)$$

operator A is strictly monotone.

Recall that $C_1 = h_m^3 / C_0$ defined in (4.16), (4.33) will be satisfied at least in case of large h_m , or equivalently, of light load.

Without this restriction, we can conclude nothing about the uniqueness from the present analysis.

4.5 CONVERGENCE OF THE PENALTY METHOD.

In Section 3.3, we introduced a penalty term in Reynolds-Hertz equations to release the constraint. Wu^[93] proved the weak convergence of the penalty solutions. Here we are able to show that, in fact, these sequences may converge strongly. We begin with a Lemma for the penalty term.

LEMMA 4.11. Define the penalty operator $\Phi : H^1_0(\Omega) \rightarrow H^{-1}$ as follows

$$\Phi(q) = q^- / \epsilon \text{ with } \epsilon > 0$$

$$q^- \equiv \min(q, 0) \equiv (q - |q|)/2 \quad (4.34)$$

Then Φ is a monotone operator:

$$\langle q_1^- - q_2^-, q_1 - q_2 \rangle \geq 0 \quad \forall q_1, q_2 \in V \quad (4.35)$$

Particularly, we have

$$\langle \Phi(q), q \rangle = \langle q^-, q \rangle / \epsilon = (\|q^-\|_{L^2})^2 / \epsilon \geq 0 \quad (4.36)$$

PROOF : For $q_1, q_2 \in H^1_0(\Omega)$, denote

$$\Omega_\alpha = \{x \in \Omega \mid q_\alpha \geq 0\}, \quad \Omega'_\alpha = \{x \in \Omega_\alpha \mid q_\alpha < 0\}, \quad \alpha = 1, 2$$

Then $q_\alpha^- = q_\alpha$ if $x \in \Omega'_\alpha$ and $q_\alpha^- = 0$ if $x \in \Omega_\alpha$. Thus

$$\begin{aligned} \langle q_1^- - q_2^-, q_1 - q_2 \rangle &= \int_\Omega [q_1^-(q_1 - q_2) - q_2^-(q_1 - q_2)] \, d\Omega \\ &= \int_{\Omega_1'} q_1(q_1 - q_2) \, d\Omega - \int_{\Omega_2'} q_2(q_1 - q_2) \, d\Omega \\ &= \int_{\Omega_1' \cap \Omega_2'} (q_1 - q_2)^2 \, d\Omega + \int_{\Omega_1' \cap \Omega_2} q_1(q_1 - q_2) \, d\Omega - \int_{\Omega_1 \cap \Omega_2'} q_2(q_1 - q_2) \, d\Omega \end{aligned}$$

In $\Omega_1' \cap \Omega_2$, $q_1 < 0$ and $q_1 - q_2 \geq 0 \Rightarrow -q_1 q_2 \geq 0$. Similarly, in $\Omega_1 \cap \Omega_2'$, $-q_1 q_2 \geq 0$.

Thus all the integrals above are nonnegative, that means $\langle q_1^- - q_2^-, q_1 - q_2 \rangle \geq 0$.

Setting $q_2 = 0$, gives (4.36). #

A direct consequence is that, as a sum of a pseudomonotone operator and a monotone operator, $A + \Phi$ is also pseudomonotone. Also,

$$|\langle \Phi(p), q \rangle| = \left| \int_\Omega (pq - |p|q) \, d\Omega \right| / 2\varepsilon \leq \|p\|_{L^2}^2 \|q\|_{L^2} / 2\varepsilon$$

This implies the boundedness of Φ . $A + \Phi$ is still bounded.

Condition (4.36) gives $\langle \Phi(q), q \rangle \geq 0, \forall q \in H^1_0(\Omega)$, so $A + \Phi$ is also coercive.

Hence according to Theorem 3.4, we conclude the following:

THEOREM 4.12. Operator $A + \Phi$ is bounded, coercive and pseudomonotone on $V = H^1_0(\Omega)$. Moreover, the penalized problem (3.27) has at least one solution. #

To show more properties of the penalty method, we next prove the convergence theorem:

THEOREM 4.13. Denote by p_ε the solution of (3.27) with $\varepsilon > 0$. As

$\varepsilon \rightarrow 0$ there exists a subsequence of $\{p_\varepsilon\}$, which converges weakly to some $p \in V$, where p is a solution of (3.24). Moreover, the subsequence also converges strongly to p in V .

PROOF: For $f \in V'$ and the solution $p_\varepsilon \in V$,

$$\langle A(p_\varepsilon), p_\varepsilon \rangle + \langle \Phi(p_\varepsilon), p_\varepsilon \rangle = \langle f, p_\varepsilon \rangle$$

By the monotonicity of the operator Φ defined in (4.36), $\langle \Phi(p_\varepsilon), p_\varepsilon \rangle \geq 0$.

We have

$$\langle A(p_\varepsilon), p_\varepsilon \rangle \leq \langle f, p_\varepsilon \rangle \leq \|f\|_{V'} \|p_\varepsilon\|_V \quad (4.37)$$

But A is coercive, and in fact from (4.14), $\langle A(p_\varepsilon), p_\varepsilon \rangle \geq C_1 \|p_\varepsilon\|_V^2 - C_2 \|p_\varepsilon\|_V$. So $\{p_\varepsilon\}$ must be bounded in V :

$$\|p_\varepsilon\|_{H^1} \leq (C_2 + \|f\|_{V'}) / C_1 \quad (4.38)$$

The space $H^1_0(\Omega)$ is reflexive. Hence, there exists a subsequence of the bounded set $\{p_\varepsilon\}$, still denoted $\{p_\varepsilon\}$, which converges weakly to some $p \in V$.

According to the Sobolev embedding theorem [1], $H^1(\Omega) \hookrightarrow L^2(\Omega)$ compactly. Hence $p_\varepsilon \rightarrow p$ strongly in $L^2(\Omega)$. $\int (|p_\varepsilon| - |p|)^2 d\Omega \leq \int |p_\varepsilon - p|^2 d\Omega \Rightarrow |p_\varepsilon| \rightarrow |p|$ strongly in L^2 . Therefore $p_\varepsilon^- \rightarrow p^-$ strongly in L^2 .

On the other hand, (4.37) implies $\|A(p_\varepsilon)\| \leq \|f\|$ and

$$(\|p_\varepsilon^-\|_{L^2})^2 = \langle p_\varepsilon^-, p_\varepsilon \rangle = \varepsilon \langle \Phi(p_\varepsilon), p_\varepsilon \rangle$$

$$= \varepsilon [\langle A(p_\varepsilon), p_\varepsilon \rangle - \langle f, p_\varepsilon \rangle] \leq 2\varepsilon \|f\|_{V'} \|p_\varepsilon\|_V \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0$$

since $\{p_\varepsilon\}$ is bounded. This means that $p_\varepsilon^- \rightarrow 0$ in an L^2 sense.

Combination of these facts gives us the conclusion that $p^- = 0$ a.e. in Ω , i.e.,

p belongs to the constraint set K .

For any $q \in K$; $q^- = 0$, then by (4.34),

$$\langle p_\epsilon^-, p_\epsilon - q \rangle = \langle p_\epsilon^- - q^-, p_\epsilon - q \rangle \geq 0$$

$$\langle A(p_\epsilon), p_\epsilon - q \rangle \leq \langle A(p_\epsilon), p_\epsilon - q \rangle + \langle p_\epsilon^-, p_\epsilon - q \rangle / \epsilon = \langle f, p_\epsilon - q \rangle$$

Then we have

$$\overline{\lim}_{\epsilon \rightarrow 0} \langle A(p_\epsilon), p_\epsilon - q \rangle \leq \langle f, p - q \rangle \quad (4.39)$$

Setting $q=p$, we obtain $\overline{\lim}_{\epsilon \rightarrow 0} \langle A(p_\epsilon), p_\epsilon - p \rangle \leq 0$.

Thus the operator A with the sequence $\{p_\epsilon\}$ satisfies condition P (3.15).

However, since A is pseudomonotone, we have from the definition (3.16),

$$\underline{\lim}_{\epsilon \rightarrow 0} \langle A(p_\epsilon), p_\epsilon - q \rangle \geq \langle A(p), p - q \rangle, \quad \forall q \in K. \text{ Recalling (4.39), we have } \langle A(p), p - q \rangle$$

$\leq \langle f, p - q \rangle$ or

$$\langle A(p) - f, q - p \rangle \geq 0 \quad \forall q \in K.$$

This indicates that p is a solution of (3.24).

Furthermore, by the use of (4.18), we have

$$\langle A(p_\epsilon) - A(p), p_\epsilon - p \rangle \geq C_1 \|p_\epsilon - p\|_1^2 - C_\eta \|p_\epsilon - p\|_{L^q} \|p_\epsilon - p\|_{H^1}$$

Since p_ϵ is the solution of penalty problem,

$$\langle A(p_\epsilon), p_\epsilon - p \rangle = \langle f - p_\epsilon^- / \epsilon, p_\epsilon - p \rangle$$

$p \in K$, $p^- = 0$ and (4.35) gives $\langle p_\epsilon^-, p_\epsilon - p \rangle = \langle p_\epsilon^- - p^-, p_\epsilon - p \rangle \geq 0$. Finally, we have

$$C_1 \|p_\epsilon - p\|_1^2 - C_2 \|p_\epsilon - p\|_{L^q} \|p_\epsilon - p\|_{H^1} \leq \langle f - A(p), p_\epsilon - p \rangle$$

$$C_1 \|p_\epsilon - p\|_1^2 \leq C_2 \|p_\epsilon - p\|_{L^q} \|p_\epsilon - p\|_{H^1} + \langle f - A(p), p_\epsilon - p \rangle \rightarrow 0. \quad \#$$

This Theorem confirms that the penalty problem (3.27) is a legitimate approximation of (3.24) and it makes sense to consider approximations of (3.27) as

a means of approximating the solution of (3.24) for ϵ sufficiently small. In particular, if one finds a convergent sequence of penalty solutions p_ϵ , one may believe the limit is a true solution to the lubrication problem and that p_ϵ is a reasonably good approximation for ϵ sufficiently small.

4.6 A REGULARITY THEOREM FOR SOLUTIONS OF THE LINE CONTACT CASE.

Because the thickness of lubricant is very small in comparison with overall dimensions of the bearing, very high pressure gradients can be expected to occur in these class of problems. These high pressure gradients are often very difficult to model analytically or numerically. The high gradients and rapid change in sign of gradients have led many investigators to speculate that the exact pressure fields exhibit a cusp-type spike near the outlet, this representing a strong singularity in pressure gradients.

In this section, the regularity of pressure solution for the line contact case is explored. It is shown that, contrary to conjectures in the literature, the pressure profile is quite smooth and is in fact a classical solution to the elastohydrodynamic lubrication problem. We begin the analysis with a Lemma revealing property of the function $h_1(p)$.

LEMMA 4.14 If $p(x) \in C^0[a,b]$, then

$$h(x) \equiv \int_a^b p(\xi) \ln |\xi - x| d\xi \in C^0[a,b] \quad (4.40)$$

If $p(x) \in C^1[a,b]$, then $h(x) \in C^1(a,b)$ and for $a < x < b$,

$$h'(x) = \int_a^b p'(\xi) \ln |\xi - x| d\xi + p(a) \ln |\xi - x| - p(b) \ln |\xi - x| \quad (4.41)$$

The existence of $h'(a+0)$ (similarly, $h'(b-0)$) depends on $p(a)$. If $p(a)=0$,

$$h'(a+0) = \int_a^b p'(\xi) \ln |\xi - x| d\xi - p(b) \ln |b-a| \quad (4.42)$$

Otherwise $h'(a+0)$ does not exist.

PROOF : (1) For $x_0 \in [a, b]$, $|x - x_0| < \delta$, $\delta > 0$, extend $p(x)$ continuously to $[a - \delta, b + \delta]$.

$$\begin{aligned}
 h(x) - h(x_0) &= \int_a^b p(\xi) [\ln |\xi - x| - \ln |\xi - x_0|] d\xi \\
 &= \int_{a-x}^{b-x} \ln |\eta| p(\eta+x) d\eta - \int_{a-x_0}^{b-x_0} \ln |\eta| p(\eta+x_0) d\eta \\
 &= \int_{a-x_0}^{b-x_0} \ln |\eta| (p(\eta+x) - p(\eta+x_0)) d\eta + \left(\int_{a-x}^{a-x_0} + \int_{b-x_0}^{b-x} \right) \ln |\eta| p(\eta+x) d\eta \\
 &= \int_a^b \ln |\eta - x_0| (p(\eta+x-x_0) - p(\eta)) d\eta \\
 &\quad + \left(\int_{a-(x-x_0)}^a + \int_b^{b-(x-x_0)} \right) \ln |\eta - x_0| p(\eta+x-x_0) d\eta \tag{4.43}
 \end{aligned}$$

p is extended continuously to $[a - \delta, b + \delta]$, so p is uniformly continuous on $[a - \delta, b + \delta]$.

Then p is uniformly bounded on $[a - \delta, b + \delta]$ and $p(\eta+x-x_0) \rightarrow p(\eta)$ uniformly with respect to η . Furthermore $\int_a^b \ln |\eta - x_0| d\eta < \infty$. Then (4.43) leads to the fact

$$\lim_{x \rightarrow x_0} h(x) = h(x_0) \quad \forall x_0 \in [a, b] \tag{4.44}$$

(2) In case $p \in C^1[a, b]$, also extend p to $[a - \delta, b + \delta]$ smoothly.

Denote $\Delta = x - x_0$, $|\Delta| < \delta$. Then (4.43) gives

$$\begin{aligned}
 \frac{h(x) - h(x_0)}{\Delta} &= \int_a^b \ln |\eta - x_0| \frac{p(\eta+\Delta) - p(\eta)}{\Delta} d\eta + \\
 &\quad + \frac{1}{\Delta} \left(\int_{a-\Delta}^a + \int_b^{b-\Delta} \right) \ln |\eta - x_0| p(\eta+\Delta) d\eta \tag{4.45}
 \end{aligned}$$

Now $p \in C^1[a - \delta, b + \delta]$, p' is uniformly continuous, then $(p(\eta+\Delta) - p(\eta))/\Delta$ must be bounded uniformly. Thus the first integral in (4.45) is dominated by a Lebesgue integrable function $\ln |\eta - x_0| p'(\eta)$ and has the limit

$$\int_a^b \ln |\eta - x_0| [p(\eta+\Delta) - p(\eta)]/\Delta d\eta \rightarrow \int_a^b \ln |\eta - x_0| p'(\eta) d\eta.$$

For $a < x_0 < b$ and small Δ , $\ln |\eta - x_0|$ is continuous in η in $(a-\Delta, a)$ and $(b, b-\Delta)$. By the mean value theorem of integral calculus, we obtain

$$\begin{aligned} & (\int_{a-\Delta}^a + \int_b^{b-\Delta}) \ln |\eta - x_0| p(\eta + \Delta) d\eta / \Delta \\ &= \ln |a - \theta_1 \Delta - x_0| p(a - \theta_1 \Delta + \Delta) - \ln |b - \theta_2 \Delta - x_0| p(b - \theta_2 \Delta + \Delta) \end{aligned}$$

$0 \leq \theta_1, \theta_2 \leq 1$. The limit as $\Delta \rightarrow 0$ is $\ln |a - x_0| p(a) - \ln |b - x_0| p(b)$. Therefore, for $a < x_0 < b$, we have (4.41).

(3) Consider the case $x_0 = a$, and $\Delta > 0$. Similarly,

$$\begin{aligned} \frac{h(a+\Delta) - h(a)}{\Delta} &= \int_a^b \ln |\eta - a| \frac{p(\eta + \Delta) - p(\eta)}{\Delta} d\eta \\ &+ \frac{1}{\Delta} (\int_{a-\Delta}^a + \int_b^{b-\Delta}) \ln |\eta - a| p(\eta + \Delta) d\eta \end{aligned}$$

Only the term $\int_{a-\Delta}^a (\dots)/\Delta$ differs from the corresponding term in (4.45). Here $\ln |\eta - a|$ has a weak singularity at $\eta = a$. For small $\Delta > 0$, $|\eta - a| < \Delta$, $\ln |\eta - a|$ has no change in sign. Then the mean value theorems of integral and differential calculus lead to

$$\begin{aligned} \int_{a-\Delta}^a \ln |\eta - a| p(\eta + \Delta) d\eta / \Delta &= p(a - \theta \Delta + \Delta) / \Delta \int_{a-\Delta}^a \ln |\eta - a| d\eta \\ &= [p(a) + (1 - \theta) \Delta p'(a + \zeta \Delta)] / \Delta \left[(\eta - a) \ln |\eta - a| - \eta \right] \Big|_{\eta=a-\Delta}^{\eta=a} \\ &= (p(a) + p'(a) O(\Delta)) (\ln |\Delta| - 1) \end{aligned}$$

Thus it is obvious if $p(a) = 0$, the last expression has limit 0 and if $p(a) \neq 0$, it diverges.

A similar conclusion is valid for $x_0 = b$.

#

Now come back to the penalized problems. For the solution $p_\epsilon \in H^1_0(a, b)$, we have $p'_\epsilon \in L^2(a, b)$. Then p'_ϵ is Lebesgue integrable over (a, b) , this implies that

p_ε is absolutely continuous. Thus, $p_\varepsilon(x) \in C^0[a,b]$.

According to the Sobolev embedding theorem, for the one-dimensional case, $H^1(\Omega) \subset C^0(\Omega)$. That means there exists a constant $C_c > 0$ such that

$$\max(q) = \|q\|_{C^0} \leq C_c \|q\|_{H^1}, \quad \forall q \in H^1$$

Recall (4.38), for the penalty solution, we have, in addition, that $p_\varepsilon(x)$ is uniformly bounded with respect to ε :

$$p_\varepsilon \in C^0(\Omega), \quad \|p_\varepsilon\|_{C^0} \leq C_c (C_2 + \|f\|_V) / C_1 \quad (4.46)$$

Lemma 4.14 results in

$$h_1(p_\varepsilon)(x) = 2/\pi E' \int_a^b \ln(\xi - x_0 / \xi - x)^2 p(\xi) d\xi \in C^0[a,b] \quad (4.47)$$

and (4.46) gives the uniform boundedness of $h_1(p_\varepsilon)(x)$.

Equation (3.27) means that for the solution p_ε , $A(p_\varepsilon) + p_\varepsilon^-/\varepsilon = f$ holds in V' . Our function f is in fact smooth enough, and now $p_\varepsilon \in C^0[a,b]$, moreover $p_\varepsilon^- \in C^0[a,b]$, thus, we have

$$A(p_\varepsilon) = -\frac{d}{dx} \left(h^3(p_\varepsilon) e^{-\alpha p_\varepsilon} \frac{dp_\varepsilon}{dx} \right) + 12\mu_0 \frac{du_0 h_1(p_\varepsilon)}{dx} \in C^0[a,b] \quad (4.48)$$

Therefore,

$$-h^3(p_\varepsilon) e^{-\alpha p_\varepsilon} dp_\varepsilon/dx + 12\mu_0 u_0 h_1(p_\varepsilon) \in C^1[a,b] \quad (4.49)$$

Here we would like to employ the expression (2.28) for film thickness $h(p_\varepsilon) = h_z + h_1(p_\varepsilon) + h_0$. Then it is straightforward that $h(p_\varepsilon) \in C^0[a,b]$, and we arrive at the following

LEMMA 4.15. The penalized problem (3.27) has solution $p_\varepsilon \in C^1[a,b]$.

#

Using Lemma 4.14 again with the conditions $p_\varepsilon(a) = p_\varepsilon(b) = 0$, we obtain

$h_1(p_\varepsilon) \in C^1[a, b]$ and

$$h_1'(p_\varepsilon) = -2/\pi E' \int_a^b \ln(\xi-x)^2 p_\varepsilon'(\xi) d\xi \quad (4.50)$$

It follows that $h(p_\varepsilon) \in C^1[a, b]$ and from (4.49)

THEOREM 4.16. The penalized problem (3.27) has a twice differentiable solution $p_\varepsilon \in C^2[a, b]$, and the corresponding film thickness $h(p_\varepsilon) \in C^1[a, b]$. #

Furthermore, we can prove

THEOREM 4.17. The sequence of one-dimensional solutions $\{p_\varepsilon\}$ is bounded in $H^2(a, b)$.

PROOF: Set $q = p_\varepsilon^-$ in (3.27). This leads to

$$\langle A(p_\varepsilon), p_\varepsilon^- \rangle + \langle p_\varepsilon^-, p_\varepsilon^- \rangle / \varepsilon = \langle f, p_\varepsilon^- \rangle$$

$$(\|p_\varepsilon^-\|_{L^2})^2 / \varepsilon = \langle p_\varepsilon^-, p_\varepsilon^- \rangle / \varepsilon = -\langle A(p_\varepsilon), p_\varepsilon^- \rangle + \langle f, p_\varepsilon^- \rangle$$

$$= \int_{(p_\varepsilon < 0)} [h^3(p_\varepsilon) e^{-\alpha p_\varepsilon} (dp_\varepsilon/dx)^2 + 12\mu_0 d(u_0 \tilde{h}_1(p_\varepsilon))/dx p_\varepsilon^-] dx + \langle f, p_\varepsilon^- \rangle$$

$$\leq \langle f, p_\varepsilon^- \rangle - \int_{(p_\varepsilon < 0)} 12\mu_0 d(u_0 \tilde{h}_1(p_\varepsilon))/dx p_\varepsilon^- dx$$

$$\leq \|f\|_{L^2} \|p_\varepsilon^-\|_{L^2} + 12\mu_0 \|d(u_0 \tilde{h}_1(p_\varepsilon))/dx\|_{L^2} \|p_\varepsilon^-\|_{L^2}.$$

From (4.4) and (4.13),

$$|h_1(p_\varepsilon)| \leq C_h \|p_\varepsilon\|_{L^4} \leq C_q C_h \|p_\varepsilon\|_{H^1} \quad (4.51)$$

and from (4.41), we have

$$|h_1'(p_\varepsilon)| \leq 4/\pi E' \|p_\varepsilon'\|_{L^2} [\int_a^b (\ln|\xi-x|)^2 d\xi]^{1/2} \leq C_h \|p_\varepsilon\|_{H^1} \quad (4.52)$$

Recalling (4.38), the boundedness of $\|p_\varepsilon\|_{H^1}$ yields the consequence that $h_1(p_\varepsilon)$ and

$h_1'(p_\epsilon)$ are uniformly bounded in $[a,b]$ with respect to ϵ . Therefore, we have

$12\mu_0\|d(u_0h_1(p_\epsilon))/dx\|_{L^2}$ bounded by some $M>0$, and $\|p_\epsilon^-\|_{L^2}/\epsilon \leq \|f\|_{L^2} + M$, or

$$\|p_\epsilon^-\|_{L^2} = O(\epsilon) \quad (4.53)$$

This indicates that $p_\epsilon^- \in L^2$. With $h(p_\epsilon) \in C^1$, we may now consider Reynolds equation in an L^2 -sense,

$$\begin{aligned} & -h^3(p_\epsilon)e^{-\alpha p_\epsilon} d^2p_\epsilon/dx^2 - d(h^3(p_\epsilon)e^{-\alpha p_\epsilon})/dx dp_\epsilon/dx \\ & + 12\mu_0 d(u_0\tilde{h}_1(p_\epsilon))/dx + p_\epsilon^-/\epsilon = f. \end{aligned}$$

with $h \geq h_m > 0$ (here we again use the expression (2.28')), $h^{-3}(p_\epsilon) \leq h_m^{-3}$.

$$\begin{aligned} d^2p_\epsilon/dx^2 = & -h^{-3}(p_\epsilon) [f + d(h^3(p_\epsilon)e^{-\alpha p_\epsilon})/dx dp_\epsilon/dx - p_\epsilon^-/\epsilon \\ & - 12\mu_0 d(u_0\tilde{h}_1(p_\epsilon))/dx] \end{aligned}$$

From (4.51-53), we know that $\|d^2p_\epsilon/dx^2\|_{L^2}$ is bounded with respect to the parameter ϵ . Then with the fact that $\{p_\epsilon\}$ is bounded in H^1 we may conclude that $\{p_\epsilon\}$ is bounded in H^2 . #

Now we are able to derive an important result.

THEOREM 4.18. The one-dimensional Reynolds-Hertz problem (3.3) has solution $p \in C^1[a,b]$.

PROOF: Theorem 4.17 shows that $\{p_\epsilon\}$ is bounded in $H^2(a,b)$ which is again a reflexive Banach space. Thus there exists a subsequence, still denoted by $\{p_\epsilon\}$, converging weakly to some p in H^2 . Applying the Sobolev embedding theorem ($n=1, m=1, p=2$ case), we have $H^2 \subset C^1$ and $H^2 \subset H^1$ compactly. So $p_\epsilon \rightarrow p$ strongly in C^1 .

Now the fact $p_\epsilon \in C^1[a,b]$ gives $p \in C^1[a,b]$. #

As a direct consequence, since $p=0$ in Ω_0 , we have

COROLLARY 4.19. For the C^1 -solution p of (3.3), Reynolds condition holds at the free boundary

$$\frac{dp}{dx} \Big|_{x=b_1} = 0 \quad (4.54) \quad \#$$

Repeating the argument in deriving $p_\epsilon \in C^2(a,b)$, we obtain the following conclusion with a simple substitution of b_1 for b .

THEOREM 4.20. The one-dimensional Reynolds-Hertz problem (3.3) has a solution in $C^2(a,b_1)$, and, in fact, this solution is a classical solution. $\#$

REMARK 1. Use of the Reynolds condition (4.54) as a criterion for locating the free boundary has been made in numerical studies of related lubrication problems for several years by many people. A physical interpretation and motivation of this condition is demonstrated in [28].

In [76], a first attempt was made to prove the continuity of p' using a maximum principle (for rigid bearings). Unfortunately there is a mistake in this work, as all derivatives concerned were from the left side (inside the contact region) of α^* (refer to [76]), so that setting $p(x,\alpha)=0$ for $x \in [\alpha,b]$ (i.e., $x \geq \alpha$) may not imply that $dp/dx|_{x=\alpha-} + \partial p/\partial \alpha|_{x=\alpha-} = 0$, which was a major condition in that work.

In addition, for elastohydrodynamic lubrication, an extremum principle may no longer exist. However we obtain the results developed here through use of a variational inequality and a penalized problem.

REMARK 2. The possibility of a singularity, a spike or cusp, in the pressure profile has long been a concern to tribologists who attempt to model elastohydrodynamic lubrication numerically. However, we have shown here, that such a spike cannot exist in the one-dimensional solution. In fact, we must have $p \in C^1[a,b]$. This result is also confirmed by our numerical experiments (see Chapter 5 of this work).

REMARK 3. Physically, the thickness of the lubricant film which is

considered as a fluid must vary smoothly, particularly at inlet before a disturbance occurs in the contact region.

Recalling Lemma 4.14 and Theorem 4.16, we established that $p_\epsilon \in C^2[a,b]$, $p_\epsilon \in C^1[a,b]$. We know, at least, that

$$g \equiv h_1'(p_\epsilon) = -4/\pi E' \int_a^b \ln |\xi-x| p_\epsilon'(\xi) d\xi \in C^1(a,b)$$

$$g' \equiv h_1''(p_\epsilon) = -4/\pi E' \left[\int_a^b \ln |\xi-x| p_\epsilon''(\xi) d\xi \right.$$

$$\left. + \ln |a-x| p_\epsilon'(a) - \ln |b-x| p_\epsilon'(b) \right]$$

That $h(p_\epsilon)''(a+0)$ exists requires $p_\epsilon'(a)=0$. This also implies $p'(a)=0$ by C^1 convergence. Although we have not yet proved this conjecture and there is not a corresponding "free boundary" at the inlet zone, it is strongly supported by the numerical experiments shown in Fig. 5.6. Of course this condition is also physically realistic.

REMARK 4. In proving Theorem 4.13, we proved that in the general case, $(\|p_\epsilon - \bar{p}\|_{L^2})^2/\epsilon$ is bounded, i.e., $\|p_\epsilon - \bar{p}\|_{L^2} = O(\sqrt{\epsilon})$. However, here we have shown that for the one-dimensional solution, (4.53), $\|p_\epsilon - \bar{p}\|_{L^2} = O(\epsilon)$.

5. FINITE ELEMENT SOLUTIONS

Several attempts at obtaining numerical solutions to lubrication problems, especially, elastohydrodynamic problems can be found in the literature. Dowson and Higginson^[28] contributed a summary on early works for line contact problems and presented a systematic procedure for the solution of such problems. In the early 1970's, finite difference methods were used for the analysis of point contact problem ^[12]. Later, Hamrock and Dowson^[41] extended the scheme used in ^[28] to point contact problems and obtained important results, based on a modified Reynolds equation with the finite difference method. Meanwhile, Oh and Rohde^[68] completed a classical finite element solution for point contact problems. The Fifth Leeds-Lyon Symposium on Tribology ^[29] was dedicated to this topic. A more recent survey can be found in ^[26].

In the lubrication community, most engineers apparently have preferred finite difference methods to finite element methods, even though finite element methods applied to lubrication problems have exhibited numerous advantages. There are relatively few papers on finite element methods for these problems compared with the other fields.

In the treatment of the free boundary, most of the previous works utilized a certain subroutine in their programs to adjust iteratively the negative part of the pressure to be zero or to modify the integral region to fit the Reynolds condition. This natural but somewhat unsophisticated procedure suffers from the fact that no theoretical predictions on the success of such adjustments or of the quality of computation are available nor indeed, are they easy to obtain, especially for two-dimensional problems.

This chapter is devoted to the presentation of finite element methods with penalty for elastohydrodynamic problems. The capability of the penalty method in locating the free boundary is found to be excellent. The results presented here, together with those of Chapter 6 provide a basis for obtaining accurate solutions of elastohydrodynamic lubrication problems with the finite element method.

5.1 BRIEF SURVEY ON FINITE ELEMENT METHODS IN TRIBOLOGY.

While the finite difference method is widely utilized by many lubrication engineers, the finite element method (FEM) is gaining in popularity in applications to various aspects of tribology. Some earlier works were summarized in the survey by Huebner^[46]. More recently, other efforts to apply FEM to EHL problems have been made, particularly, with the free boundary taken into account. Figure 5.1 shows the representative literatures of the FEM applied to various aspects of laminar flow of lubricants.

5.2 FINITE ELEMENT FORMULATION.

Following more or less standard procedures (referring to [7,52]), we develop a finite element approximation to the penalized problem (3.27).

A nice domain Ω , large enough to include the effective contact region, is partitioned into a mesh of finite elements:

$$\Omega = \bigcup_e \Omega_e, \quad \Omega_e \cap \Omega_f = \emptyset \quad \text{if } e \neq f$$

associated with a set of piecewise polynomial shape functions $\{\Phi_n\}$. Denote

$$V_h = \{\text{linear combinations of } \{\Phi_n\} \text{ over } \Omega\} \quad (5.1)$$

Then we interpolate the approximate solution p_ϵ^h as

$$p_\epsilon^h = \sum_{n=1}^N p^n \Phi_n(x)$$

$$p_\epsilon^h|_e = \sum_{n=1}^{N_e} p_e^n \Phi_{ne}(x) \quad (5.2)$$

Here the index (e) indicates quantities defined on the element level. The index (h), the mesh parameter, expresses conventionally the finite element approximation; $\{\Phi_n\}$ are global shape functions; p^n and p_e^n are nodal values of p_ϵ .

Now we have an approximate problem:

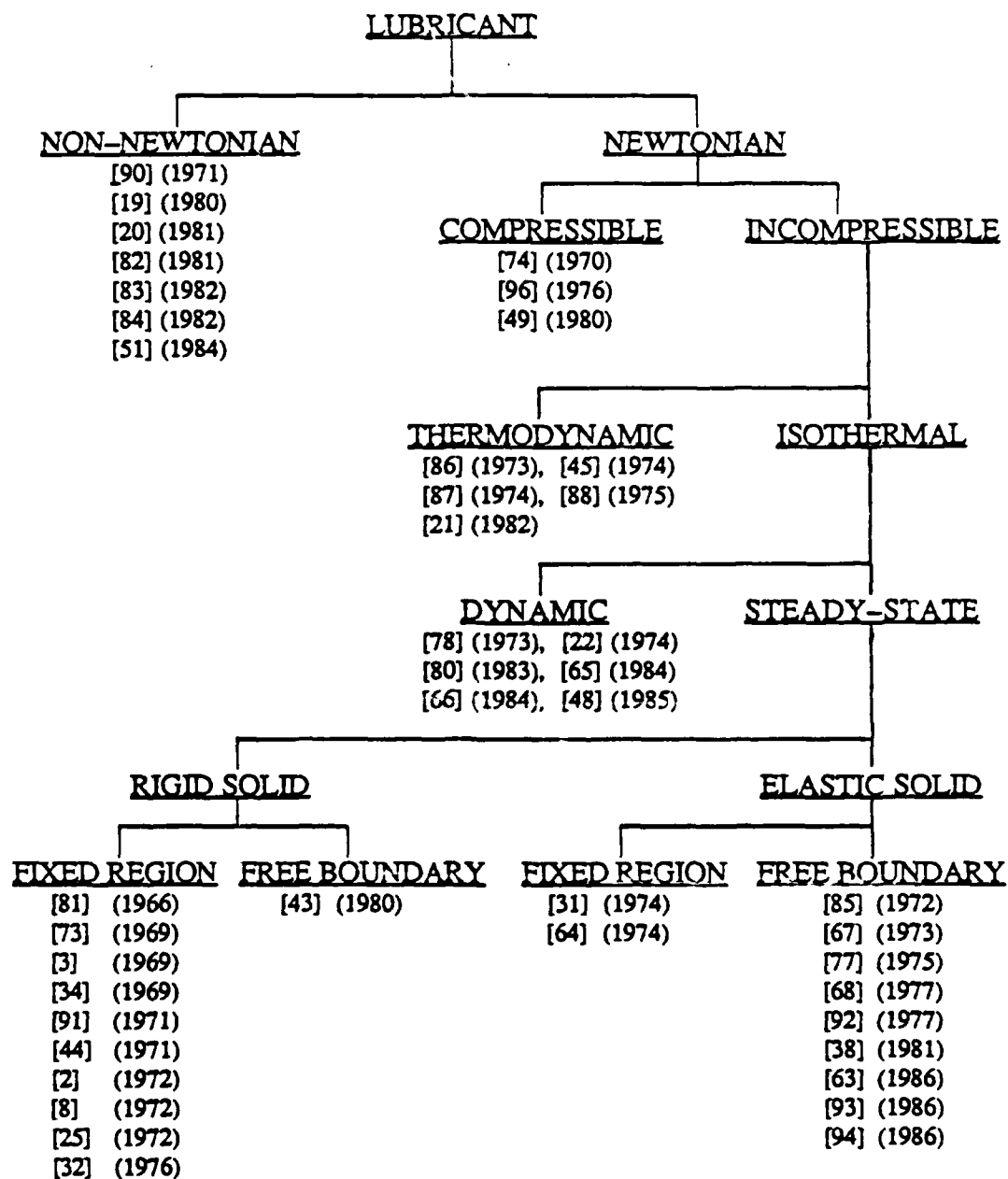


Fig. 5.1 Survey on FEM Applied to Lubrication

(P_ϵ^h) : Find $p_\epsilon^h \in V^h \subset H_0^1(\Omega)$ such that

$$\langle A(p_\epsilon^h), q^h \rangle + \langle (p_\epsilon^h)^-, q^h \rangle / \epsilon = \langle f, q^h \rangle \quad \forall q^h \in V^h \quad (5.3)$$

Introduction of (5.2) into (5.3) leads to a system of nonlinear equations in the nodal values p^n as follows

$$\begin{aligned} & \sum_{j=1}^N \left[\int_{\Omega} (h^3(p_\epsilon^h) \nabla \Phi_i \cdot \nabla \Phi_j + \frac{\Delta(p)}{\epsilon} \Phi_i \Phi_j) d\Omega \right] p^j - \int_{\Omega} 12\mu_0 u_0 \tilde{h}_1(p_\epsilon^h) \frac{\partial \Phi_i}{\partial x} d\Omega \\ & = \int_{\Omega} 12\mu_0 u_0 \tilde{h}_2 \frac{\partial \Phi_i}{\partial x} d\Omega, \quad i=1, \dots, N \end{aligned} \quad (5.4)$$

Here

$$\Delta(p) = \begin{cases} 0 & \text{if } p_\epsilon^h \geq 0 \\ -1 & \text{if } p_\epsilon^h < 0 \end{cases} \quad (5.5)$$

Note that (5.4) contains $\{p^j\}$ implicitly in the left hand side, since h , \tilde{h}_1 and $\Delta(p)$ depend on p_ϵ^h .

For the film thickness we also use interpolations

$$\begin{cases} h^h = \sum h^k \Phi_k \\ h^k = \tilde{h}_1^k + \tilde{h}_2^k \\ \tilde{h}_\alpha^k = \tilde{h}_\alpha | \text{ node } k, \alpha = 1, 2 \end{cases} \quad (5.6)$$

5.3 ITERATIVE SCHEMES.

For the nonlinear discretized equations (5.4), several iterative methods can be chosen in the numerical processes. The simplest is the so-called direct iteration. A pressure distribution $(p)_k$ is assumed first, e.g., $(p)_k \equiv 0$, or a solution for lighter load, then calculate $h_1(p_k)$ and h . In the next iteration, the equations for $(p)_{k+1}$ are solved with h_1 and h evaluated from $(p)_k$. Repeat these steps until the difference between the successive solutions is smaller than some given tolerance.

Customarily, we may formulate (5.4) as follows

$$\begin{aligned} & \sum_{j=1}^N \int_{\Omega} (h^3(p_k) e^{-\alpha p_k} \nabla \Phi_i \cdot \nabla \Phi_j + \Delta(p_k) \Phi_i \Phi_j / \epsilon) d\Omega \quad (p^j)_{k+1} \\ & = \int_{\Omega} 12\mu_0 u_0 h(p_k) \partial \Phi_i / \partial x \, d\Omega \quad i=1, \dots, N \end{aligned} \quad (5.7)$$

This formulation results in a banded, sparse, symmetric matrix form which is easy to be solved. However, such a direct iteration can only work for very light loads. For heavier loads, we would like to utilize Newton-Raphson iteration. If we denote by $G_i = 0$, $i=1, \dots, N$, for the equations (5.4) with G_i defined by

$$\begin{aligned} G_i &= \sum_{j=1}^N \int_{\Omega} (h^3 e^{-\alpha p} \nabla \Phi_i \cdot \nabla \Phi_j + \Delta(p) \Phi_i \Phi_j / \epsilon) d\Omega \, p^j \\ & \quad - \int_{\Omega} 12\mu_0 u_0 h \partial \Phi_i / \partial x \, d\Omega \end{aligned} \quad (5.8)$$

then the Newton-Raphson iteration suggests the formulation

$$\begin{aligned} & \sum_{j=1}^N \frac{\partial G_i}{\partial p^j} \bigg|_{p_k} \Delta p^j = -G_i \bigg|_{p_k} \\ & (p_{k+1})^i = (p_k)^i + \Delta p^i \quad i=1, \dots, N \end{aligned} \quad (5.9)$$

Here we have

$$\begin{aligned} \partial G_i / \partial p^j &= \int_{\Omega} (h^3(p) e^{-\alpha p} \nabla \Phi_i \cdot \nabla \Phi_j + \Delta(p) \Phi_i \Phi_j / \epsilon) \\ & \quad - \int_{\Omega} \alpha \Phi_j h^3 e^{-\alpha p} \nabla \Phi_i \cdot \nabla p \, d\Omega \\ & \quad + \int_{\Omega} (3h^2 e^{-\alpha p} \nabla \Phi_i \cdot \nabla p - 12\mu_0 u_0 \partial \Phi_i / \partial x) \partial h / \partial p^j \, d\Omega \end{aligned} \quad (5.10)$$

Recalling (2.28), we have $\partial h / \partial p^j = \partial h_1(p) / \partial p^j$, and from the fact that $h_1(p)$ is linear

in $p = \sum \Phi_j p^j$,

$$\partial h / \partial p^j = h_1(\Phi_j) \quad (5.11)$$

Generally, (5.7) can work for most cases with a large range of loads.

However, this formulation leads to an unsymmetric full matrix which is sometimes inconvenient for two-dimensional problems. Motivated by the ideas of various iterative techniques in [40,95], we may modify (5.7) and (5.9) by splitting $h_1(p)$. In fact, in finite element methods, we perform the integration on element level,

$$\int_{\Omega} (\dots) d\Omega = \sum_{e=1}^E \int_{\Omega_e} (\dots) d\Omega_e$$

In evaluating $(h_1(p))^i$, we may set two parts

$$\begin{aligned} h_1^i(p) &= h_1(p)|_{\text{node } i} = h_{10}^i + h_{11}^i \\ h_{10}^i &= \sum_{e_0} \int_{\Omega_{e_0}} C_E K(x_i, \xi) p(\xi) d\Omega \\ h_{11}^i &= \sum_{e_1} \int_{\Omega_{e_1}} C_E K(x_i, \xi) p(\xi) d\Omega = \sum_{j=1}^N H_{ij} p^j \\ H_{ij} &= \sum_{e_1} \int_{\Omega_{e_1}} C_E K(x_i, \xi) \Phi_j(\xi) d\Omega \end{aligned} \quad (5.12)$$

where e_1 denotes the elements, to which node x_i belongs, and e_0 for those elements which do not contain x_i . Thus, we may reformulate (5.7) into the H-split form,

$$\begin{aligned} \sum_{j=1}^N \left[\int_{\Omega} (h^3(p_k) e^{-\alpha p_k} \nabla \Phi_i \cdot \nabla \Phi_j + \frac{\Delta(p_k)}{\epsilon} \Phi_i \Phi_j - 12\mu_0 u_0 \frac{\partial \Phi_i}{\partial x} \sum_t H_{tj} \Phi_t) d\Omega \right] p^j \\ = \int_{\Omega} 12\mu_0 u_0 \frac{\partial \Phi_i}{\partial x} (h_2 + \sum_t h_{10}^t \Phi_t) d\Omega, \quad i=1, \dots, N \end{aligned} \quad (5.13)$$

System (5.13) is still in banded and sparse form like (5.7).

Also, we may modify Newton-Raphson iteration based on H-split formulation and use an inner iterative process in Newton-Raphson iteration which results in a banded sparse matrix like that formulated for direct iteration in (5.13). Denote $G_{ij} = \partial G_i / \partial p^j$, split $G_{ij} = G^1_{ij} + G^2_{ij}$ and make an inner iteration for (5.9)

$$\sum G^1_{ij}|_{(k)} \Delta p^j_{s+1} = -G_i|_{(k)} - \sum G^2_{ij}|_{(k)} \Delta p^j_s, \quad s=0, 1, \dots \quad (5.14)$$

where

$$G^1_{ij} = \int_{\Omega} [h^3 e^{-\alpha p} \nabla \Phi_i \cdot \nabla \Phi_j + \Delta(p) \Phi_i \Phi_j / \epsilon - \alpha \Phi_j h^3 e^{-\alpha p} \nabla \Phi_i \cdot \nabla p + (3h^2 e^{-\alpha p} \nabla \Phi_i \cdot \nabla p - 12\mu_0 u_0 \partial \Phi_i / \partial x) h_{11}(\Phi_j)] d\Omega \quad (5.15)$$

$$G^2_{ij} = \int_{\Omega} (3h^2 e^{-\alpha p} \nabla \Phi_i \cdot \nabla p - 12\mu_0 u_0 \partial \Phi_i / \partial x) h_{10}(\Omega_j) d\Omega \quad (5.16)$$

In our numerical experiments on point contact problems, formulation (5.13) can lead to the convergent solutions for quite a wide range of loads, particularly when using linear elements, although it is essentially a kind of direct iteration. Thus it can save substantial computer storage and time.

5.4 NUMERICAL RESULTS.

Nondimensional parameters are convenient for showing the basic behavior of the numerical results. Traditionally, for line contact problems, we use the following set of nondimensional parameters, with R , the radius of roller and E' , the effective elastic modulus:

$$P = p/E'; \quad G = \alpha E'; \quad U = \mu_0 u_0 / E'R; \quad W = w/E'R; \\ X = x/R; \quad H = h/R; \quad \epsilon = \epsilon R$$

where $w = \int p d\Omega$ is the load. For point contact, the only change made is for the load parameter $W = w/E'R^2$. Equation (3.3), in nondimensional form, becomes

$$-\nabla \cdot (H^2 e^{-GP} \nabla P) + P/\epsilon = -12\partial(UH)/\partial X \quad (5.17)$$

We perform numerical experiments on the penalized problems, for both line contact and point contact cases. For line contact, we use the parameters $U=10^{-11}$, $G=2500$, for point contact, we use $U=0.1683 \times 10^{-11}$, $G=4522$. Then we compute results for different values of H_0 or H_m corresponding to different loads.

Simply speaking, the penalty method is seen to work very well in locating the free boundary. Tables 5.1 and 5.2 show the sample tests which describe the behavior of pressure profile near the outlet. Only very small values of negative pressure appear, and when we reduce ϵ , they also reduce the magnitude in a stable manner, meanwhile with the positive part almost unchanged.

Table 5.1 Pressure near outlet (line contact), p*E6

(a) Linear elements solution ($H_0=0.5E-4$)

nodal x-coordinate	.00125	.00250	.00375	.00500	.00625	.00750	.00875	.01
=E4	16.382	7.3595	1.8344	-.4257E-2	-.8003E-2	-.8024E-2	-.1345E-1	0.
E3	16.384	7.3626	1.8380	-.4218E-3	-.8052E-3	-.7949E-3	-.1363E-2	0.
E2	16.384	7.4629	1.8383	-.4251E-4	-.8127E-4	-.8010E-4	-.1376E-3	0.
10	16.384	7.3629	1.8384	-.4807E-5	-.9193E-5	-.9056E-5	-.1557E-4	0.
1	16.384	7.3629	1.8384	-.4364E-6	-.8357E-6	-.8208E-6	-.1422E-5	0.
.1	16.384	7.3629	1.8384	-.4781E-7	-.9264E-7	-.8850E-7	-.1639E-6	0.
.01	16.384	7.3629	1.8384	-.4248E-8	-.8589E-8	-.7396E-8	-.1724E-7	0.
E-3	16.384	7.3629	1.8384	-.4891E-9	-.6847E-9	-.1249E-8	-.2894E-9	0.
E-4	16.384	7.3629	1.8384	-.749E-10	-.1080E-9	-.1872E-9	-.208E-10	0.
E-5	16.384	7.3629	1.8384	-.784E-11	-.115E-10	-.193E-10	-.465E-12	0.

(b) Quadratic elements solution ($H_0=0.5E-4$)

nodal x-coordinate	.00125	.00250	.00375	.00500	.00625	.00750	.00875	.01
=E4	16.468	7.4307	1.8818	-.7730E-2	-.9867E-2	-.7584E-2	-.1620E-1	0.
E3	16.473	7.4364	1.8883	-.7602E-3	-.9939E-3	-.7388E-3	-.1642E-2	0.
E2	16.473	7.4369	1.8889	-.7592E-4	-.9953E-4	-.7371E-4	-.1645E-3	0.
10	16.473	7.4370	1.8889	-.7588E-5	-.9954E-5	-.7363E-5	-.1646E-4	0.
1	16.473	7.4370	1.8890	-.7565E-6	-.9966E-6	-.7315E-6	-.1652E-5	0.
.1	16.473	7.4370	1.8890	-.7328E-7	-.1009E-6	-.6800E-7	-.1714E-6	0.
.01	16.473	7.4370	1.8890	-.5912E-8	-.1085E-7	-.3730E-8	-.2162E-7	0.
E-3	16.473	7.4370	1.8890	-.1242E-8	-.7476E-9	-.1778E-8	-.4231E-9	0.
E-4	16.473	7.4370	1.8890	-.1176E-9	-.774E-10	-.1641E-9	-.565E-10	0.
E-5	16.473	7.4370	1.8890	-.117E-10	-.777E-11	-.163E-10	-.582E-12	0.

Table 5.2 Pressure near outlet (point contact), P*E4

(a) Bilinear elements solution ($H_0=0.5E-5$)

nodal x-coordinate y=0	-.001	0.0	.001	.002	.003	.004	.005
-E4	.1576E1	.6840	.3714E-2	-.4027E-5	-.5147E-5	-.1062E-4	0.
E3	.1577E1	.6875	.9006E-4	-.1074E-5	-.4218E-6	-.1252E-5	0.
E2	.1577E1	.6886	.5378E-2	-.1211E-6	-.3886E-7	-.1267E-6	0.
10	.1577E1	.6887	.5707E-2	-.1645E-7	-.5456E-8	-.1751E-7	0.
1	.1577E1	.6887	.5710E-2	-.1199E-8	-.7071E-9	-.1792E-8	0.
.1	.1578E1	.6886	-.1720E-3	-.4348E-10	-.5307E-10	-.1107E-9	0.
.01	.1578E1	.6886	-.1074E-4	-.1332E-10	-.1679E-10	-.3480E-10	0.
E-3	.1578E1	.6887	-.6717E-6	-.1893E-11	-.2396E-11	-.4963E-11	0.
E-4	.1578E1	.6887	-.4198E-7	-.2243E-12	-.2844E-12	-.5889E-12	0.
E-5	.1578E1	.6887	-.2624E-8	-.2426E-13	-.3125E-13	-.6469E-13	0.

(b) Biquadratic elements solution ($H_0=0.5E-5$)

nodal x-coordinate y=.001	-.001	0.0	.001	.002	.003	.004	.005
-E4	.1465E1	.6038	.2227E-1	-.1585E-5	.2811E-2	.5220E-3	0.
E3	.1465E1	.6038	.2228E-1	-.1584E-6	.2828E-2	.6079E-3	0.
E2	.1465E1	.6038	.2228E-1	-.1582E-7	.2829E-2	.6166E-3	0.
10	.1465E1	.6038	.2228E-1	-.1585E-8	.2829E-2	.6174E-3	0.
1	.1465E1	.6038	.2228E-1	-.2045E-9	.2829E-2	.6175E-3	0.
.1	.1465E1	.6038	.2230E-1	-.2982E-10	.2833E-2	.6182E-3	0.
.01	.1465E1	.6040	.2262E-1	-.3340E-11	.2873E-2	.6270E-3	0.
E-3	.1466E1	.6059	.2574E-1	-.3164E-12	.3270E-2	.7137E-3	0.

5.4.1 FEM SOLUTIONS FOR LINE CONTACT PROBLEMS.

The parameters are chosen to be the same as used in [38]. Here we use $\|p_{k+1} - p_k\|_H / \|p_k\|_H < 10^{-3}$ as the convergence criterion for the iterative procedures.

In Fig. 5.2, the pressure distributions for different H_m 's, the minimum film thickness (referring (2.28')), computed with quadratic elements are shown. Figure 5.3 shows the computed load parameter W versus H_m . Figure 5.4 shows a film shape with the parameter $H_m = 1.375 \times 10^{-5}$. These results are in excellent agreement with those presented in [38], obtained with a boundary adjustment scheme. Figure 5.5 shows the linear-element solutions with the control on the parameter H_0 (referring (2.28)), which exhibit the same behavior as those calculated with H_m .

It is observed that, even under heavy loads, the pressure profiles are quite smooth, especially, the second peak is almost always located at the inner node of a quadratic element. Therefore it is in fact a smooth peak. Also, it is found that at the outlet, the free boundary condition $dp/dx = 0$ seems to hold with high accuracy. These facts confirm the regularity theorem proved in Chapter 4. Now, with the penalty method, we need not care about the location of the free boundary, contrary to the traditional boundary-adjustment iteration methods.

It is interesting to observe the pressure variation near the inlet, shown in Fig. 5.6, where $dp/dx = 0$ seems to hold perfectly. The vanishing pressure gradient at the inlet, while seldom mentioned in tribology literature, turns out to be not only a necessary condition, from the physical point of view proposed in Section 4.6, but also a natural mathematical consequence of our particular formulation of this problem.

5.4.2 FEM SOLUTIONS FOR POINT CONTACT PROBLEMS.

By the symmetry of the geometry, only the upper half of the domain Ω , the part $y \geq 0$, is needed. And a boundary condition, $\partial p / \partial y |_{y=0}$, is imposed for this situation, which does not have additional contribution to the finite element formulation. But we should include the contribution of the other half part to the

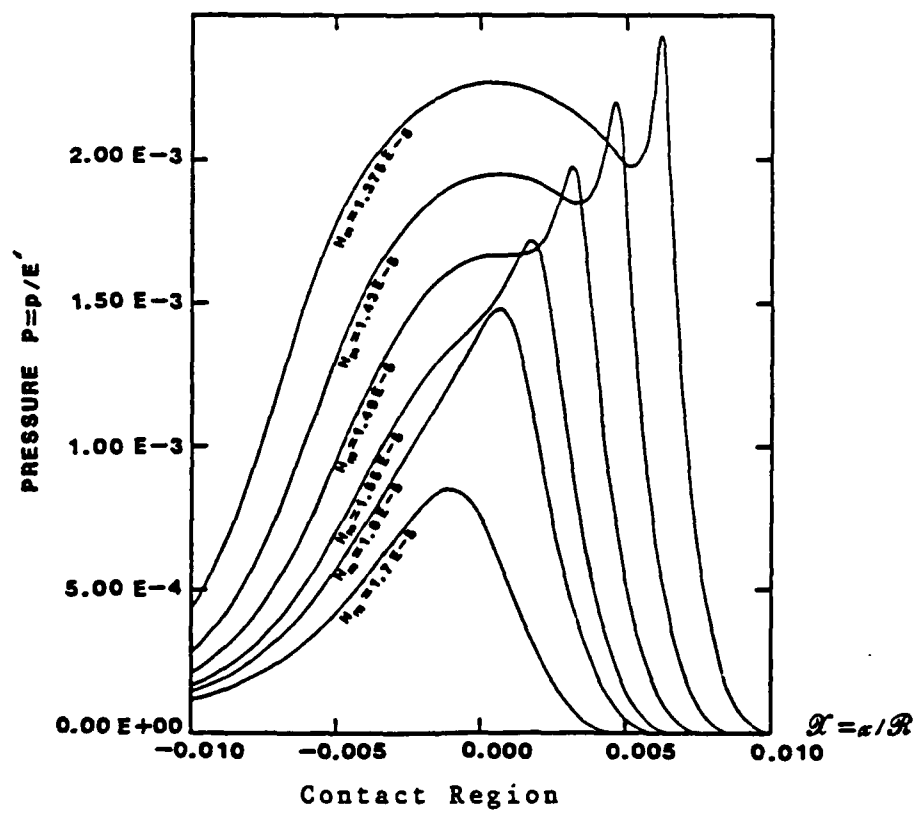


Fig. 5.2 Pressure Distributions for Line Contact (Quadratic Elements)

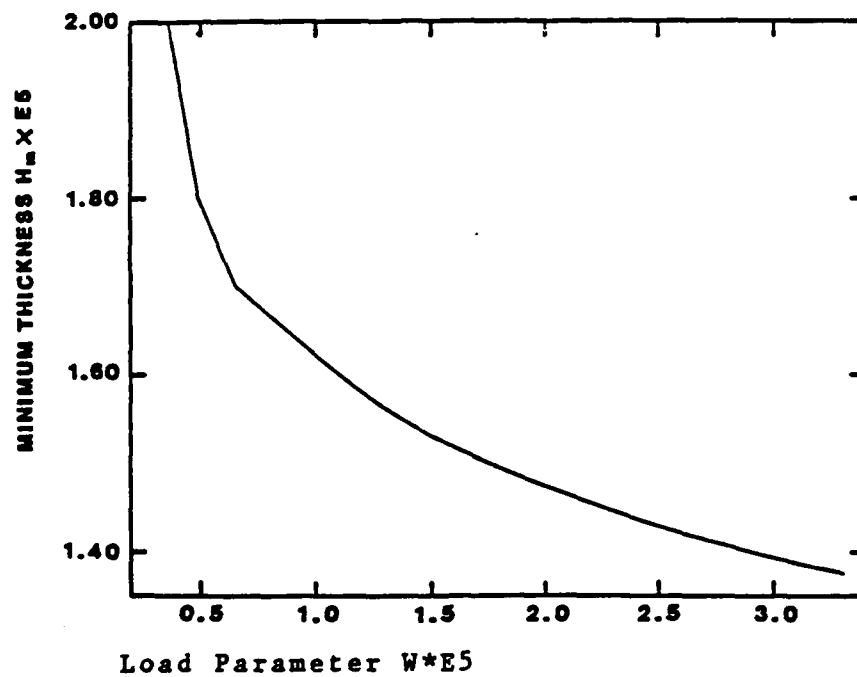


Fig. 5.3 Load Versus Minimum Film Thickness (Line Contact)

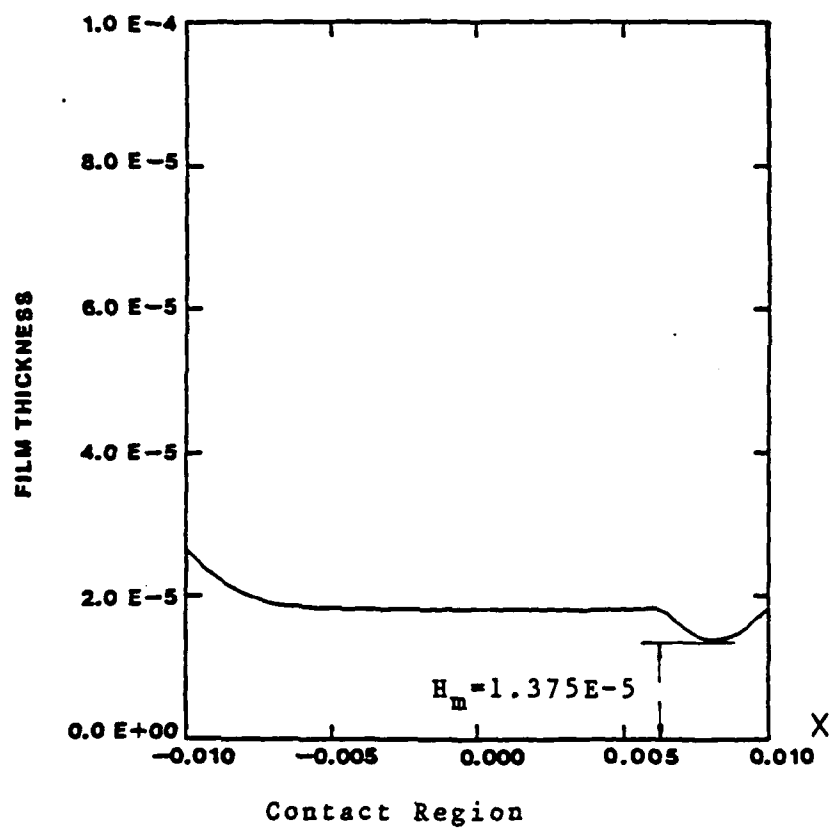


Fig. 5.4 Lubricant Film Shape (Line Contact)

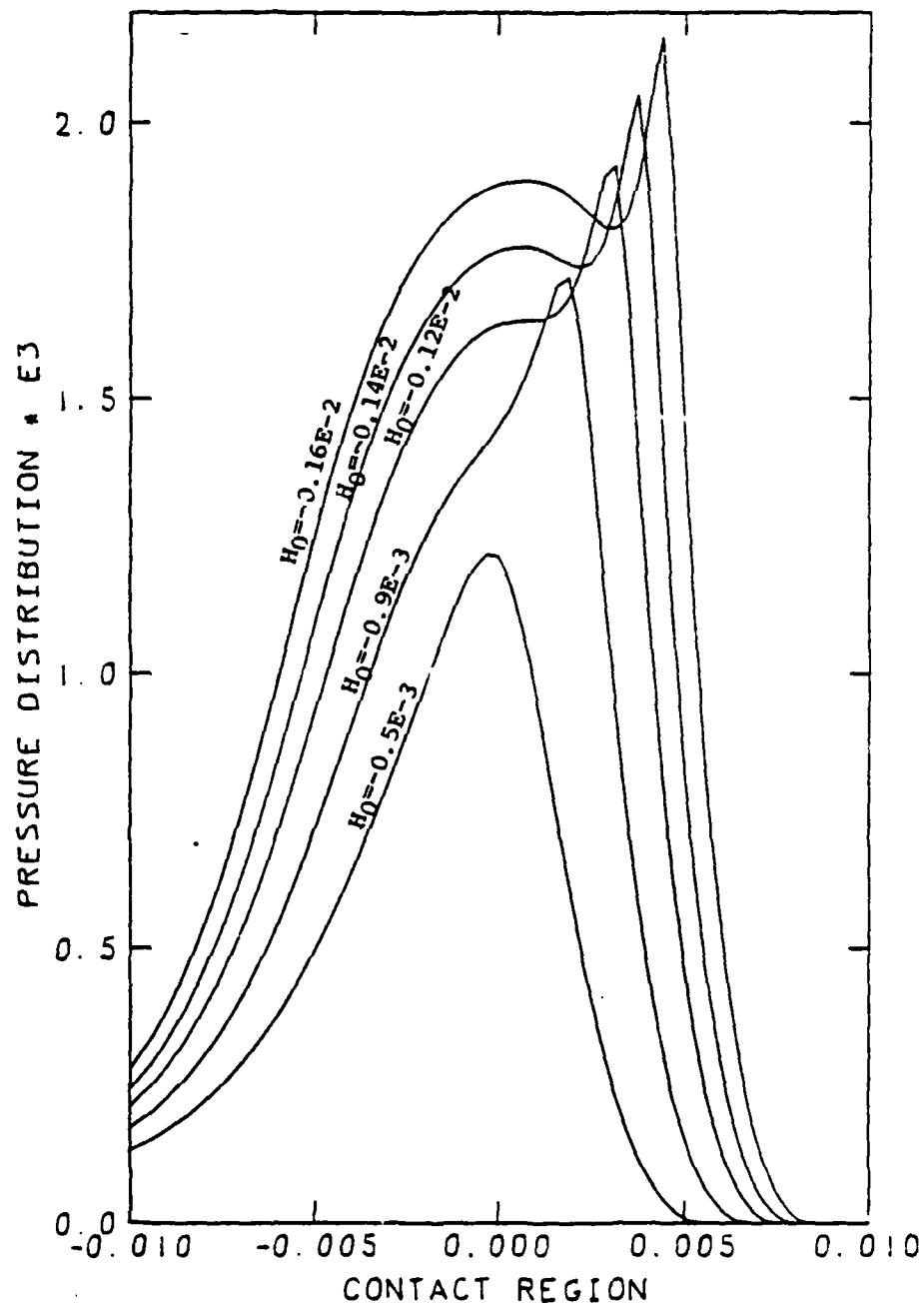


Fig. 5.5 Pressure Distributions for Line Contact (Linear Elements)

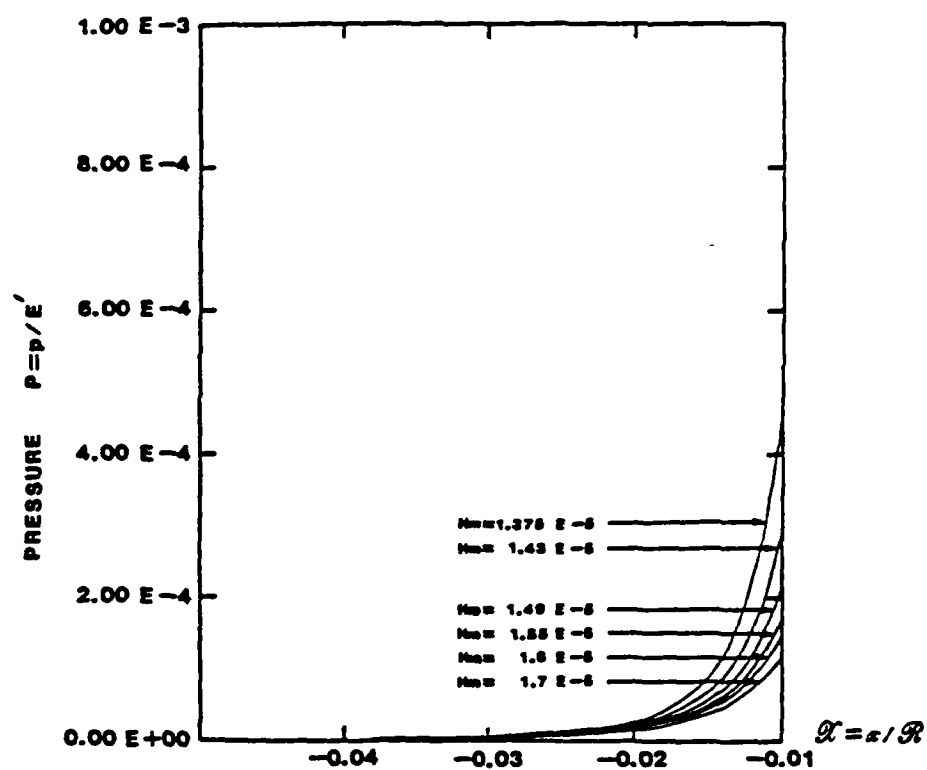


Fig. 5.6 Pressure Distributions at Inlet Zone (Line Contact)

deformation, so we have

$$h_1(p) = C_E \int p(\xi, \eta) [1/\sqrt{(x-\xi)^2 + (y-\eta)^2} + 1/\sqrt{(x-\xi)^2 + (y+\eta)^2}] d\Omega$$

As mentioned before, the penalty method works excellently in locating the free boundary. In fact, the negative part of pressure is very small, and may become even smaller when the penalty parameter ϵ is reduced. As a matter of fact, for two-dimensional problems, the Reynolds condition becomes difficult to verify, and the boundary adjustment is much more difficult than that in one-dimensional case. However, we still need not care about the location of the free boundary, an advantageous property of the penalty method.

The result in Fig. 5.7 shows a part of the pressure distribution computed with the bilinear elements. Figure 5.8 shows the pressure contours of the solution. Figure 5.9 presents the enlarged part of the pressure contours near the contact center. Figure 5.10 shows a part of film shape near the contact center. Figure 5.11 to 5.13 show a solution computed with biquadratic elements. And Fig. 5.14 shows a part of the film shape.

It is observed that the Reynolds flux condition seems to hold once again at the outlet.

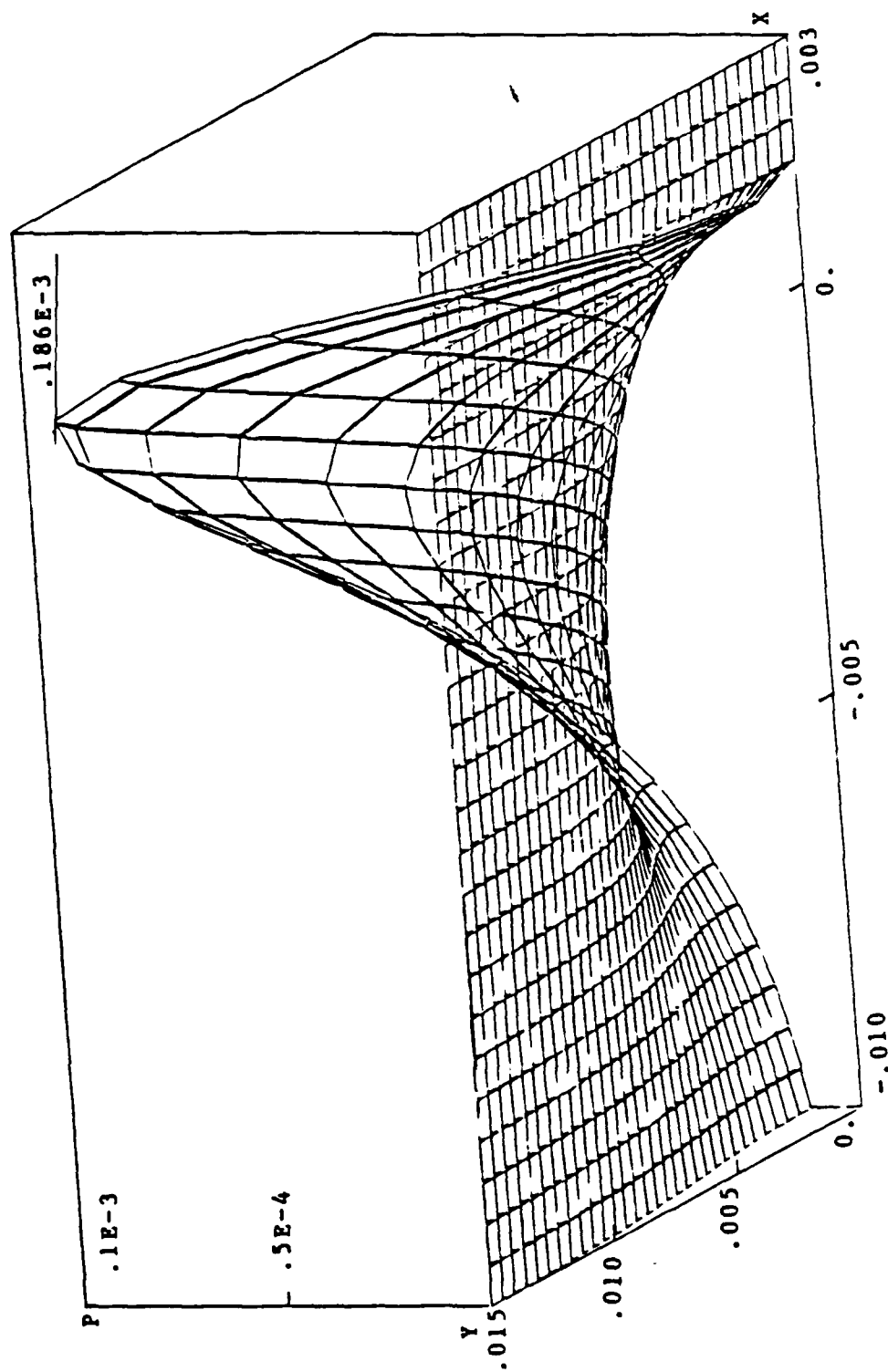


Fig. 5.7 Pressure Distribution (2-D Bilinear Elements, $H_0 = 0.5E-5$)

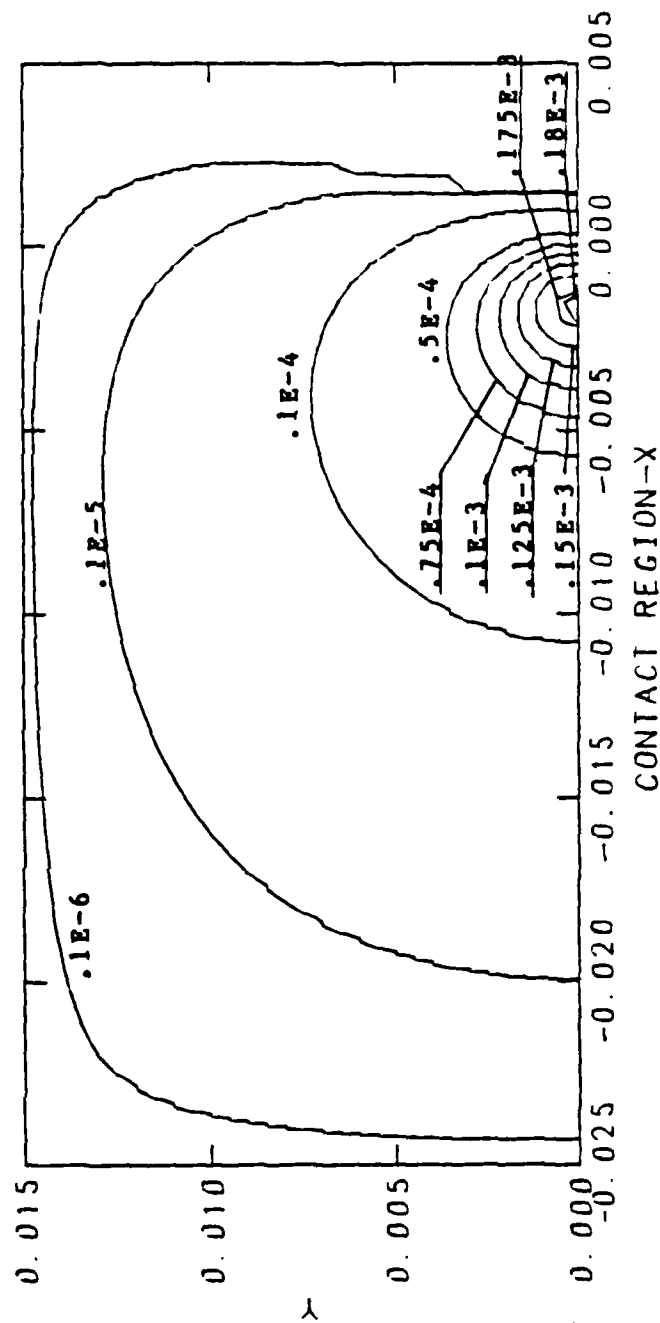


Fig. 5.8 Pressure Contours (2-D Bilinear Elements, $H_0=0.5E-5$)

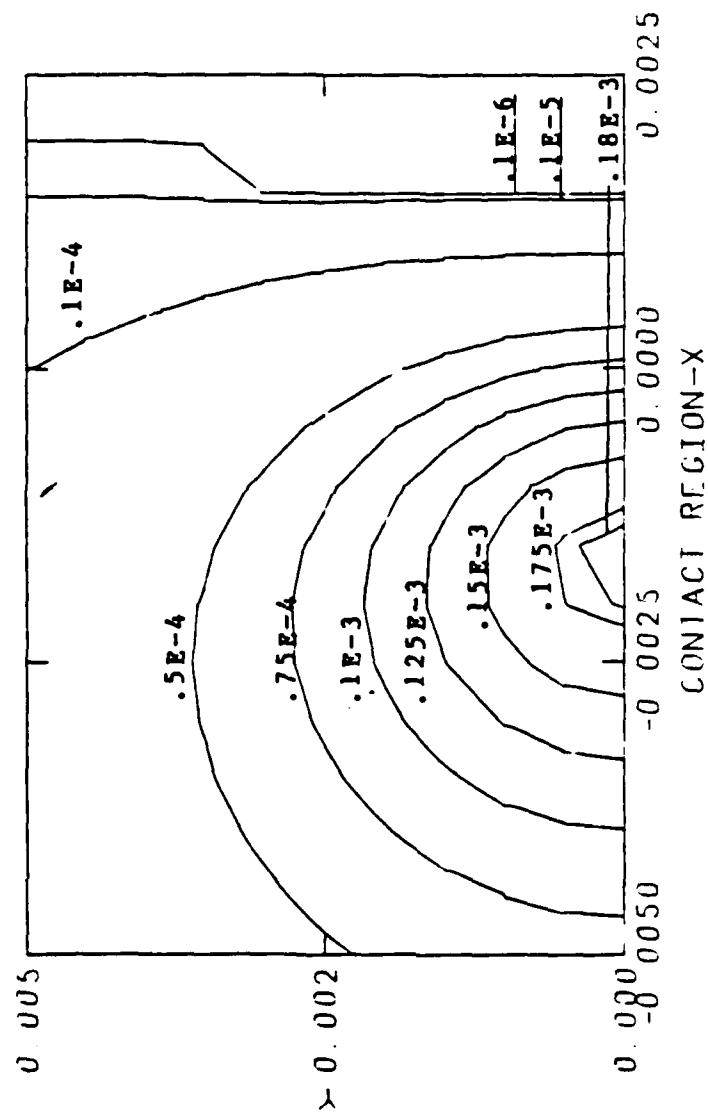


Fig. 5.9 Enlarged Part of Pressure Contours of Fig. 5.8

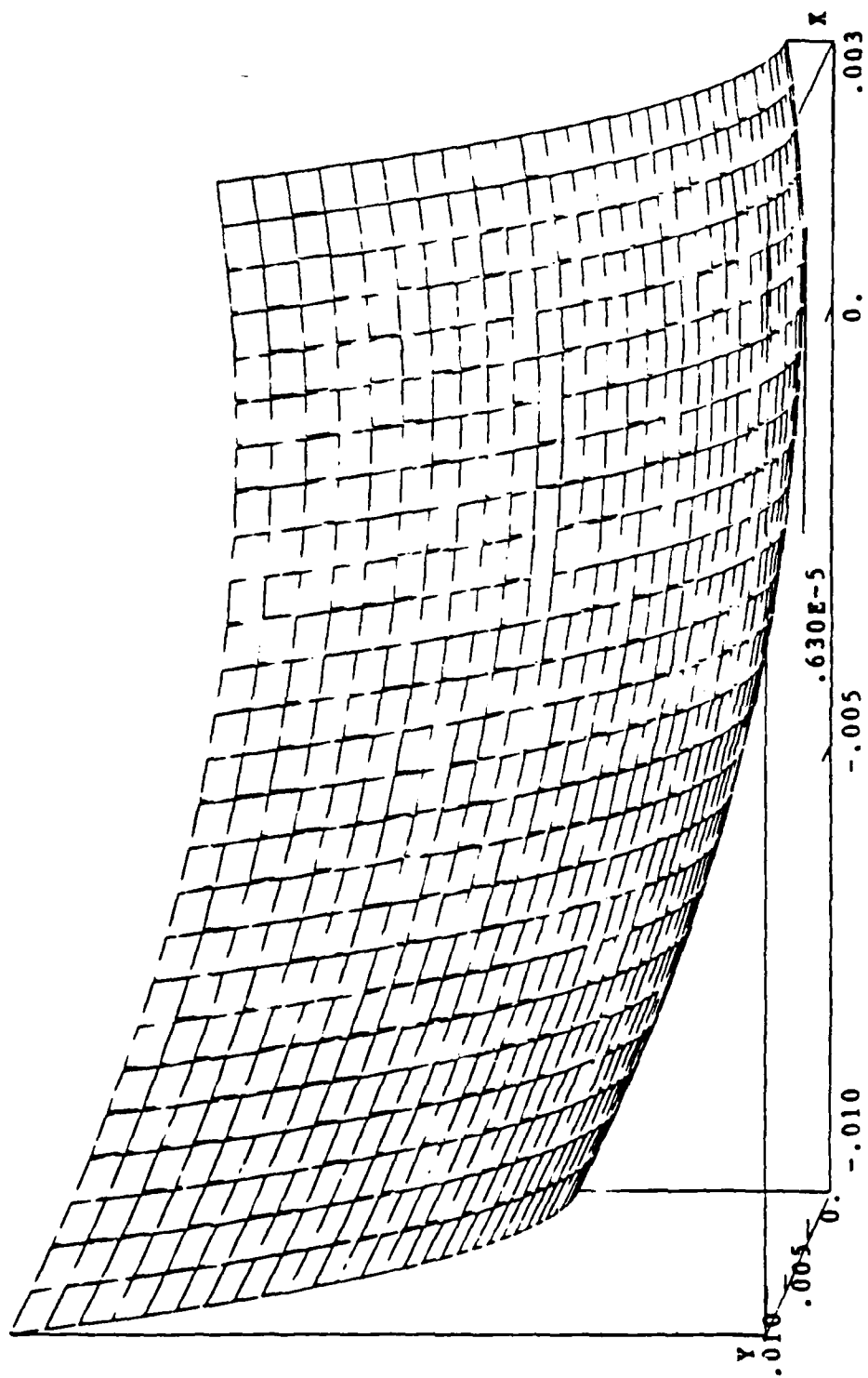


Fig. 5.10 Film Shape Near Contact Center (2-D Bilinear Elements)

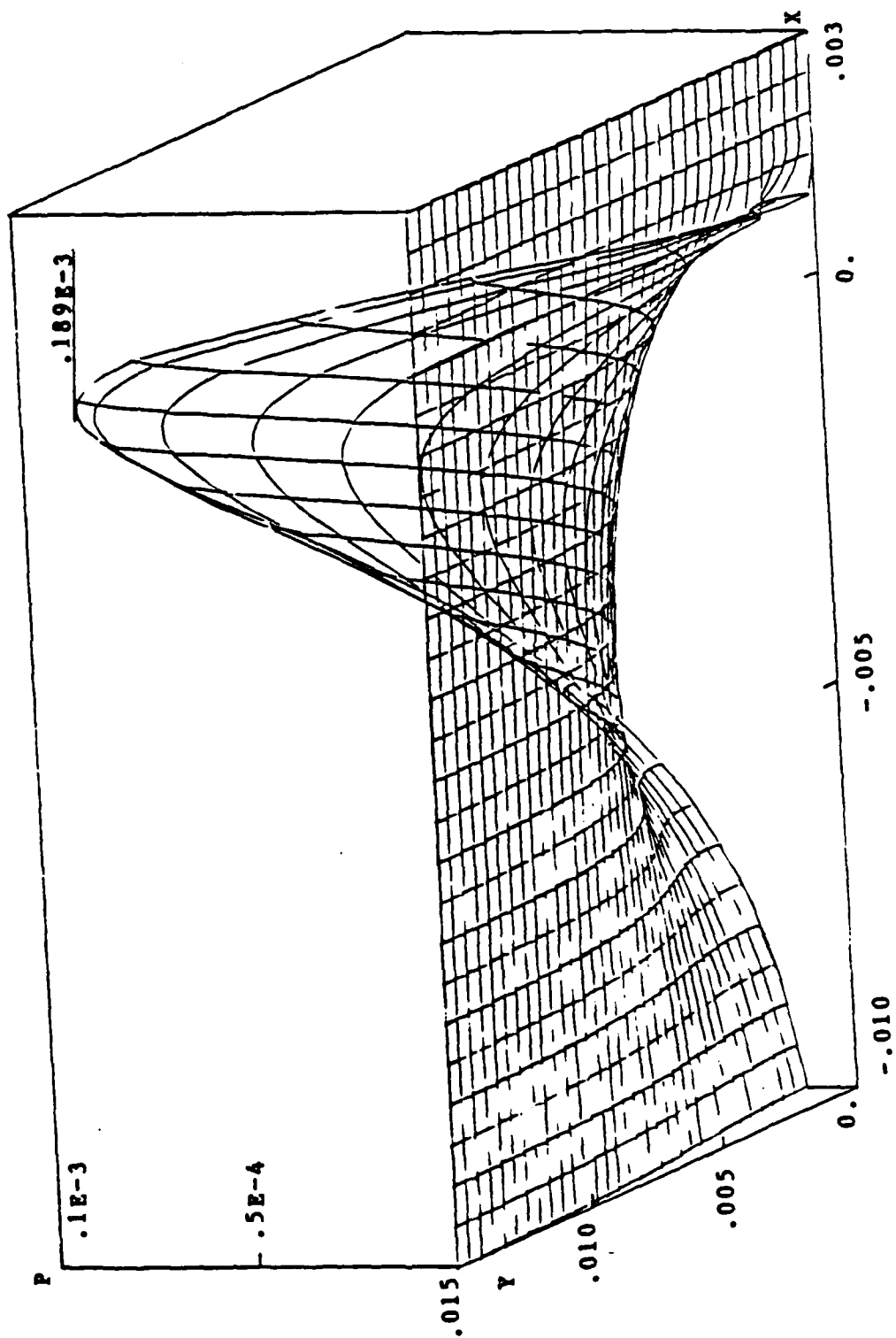


Fig. 5.11 Pressure Distribution (2-D Biquadratic Elements, $H_0=0.5E-5$)

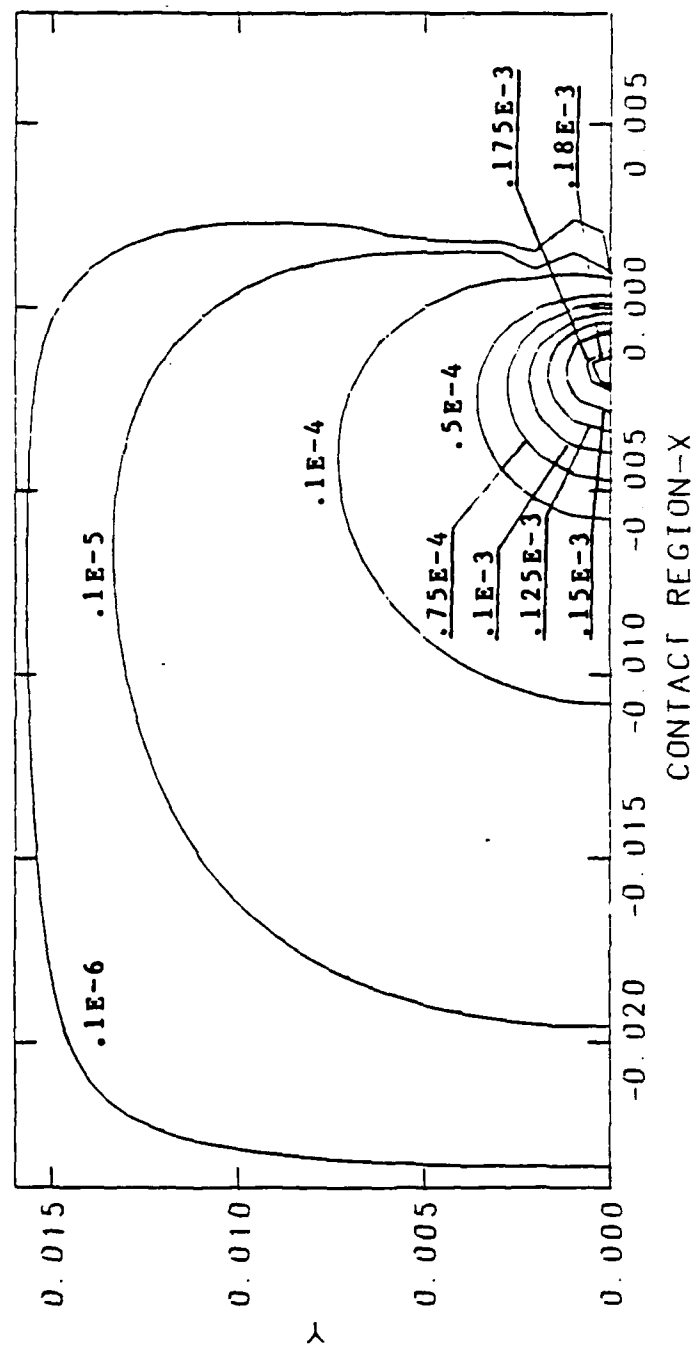


Fig. 5.12 Pressure Contours (2-D Biquadratic Elements, $H_0=0.5E-5$)

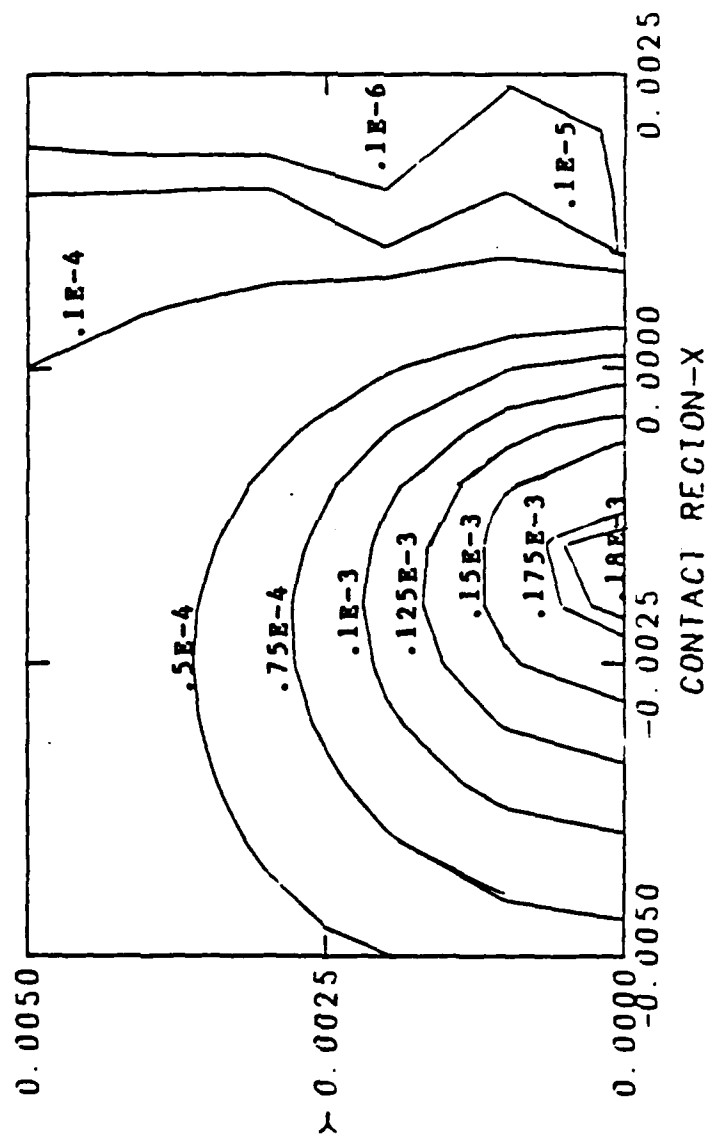


Fig. 5.13 Enlarged Part of Pressure Contours of Fig. 5.12

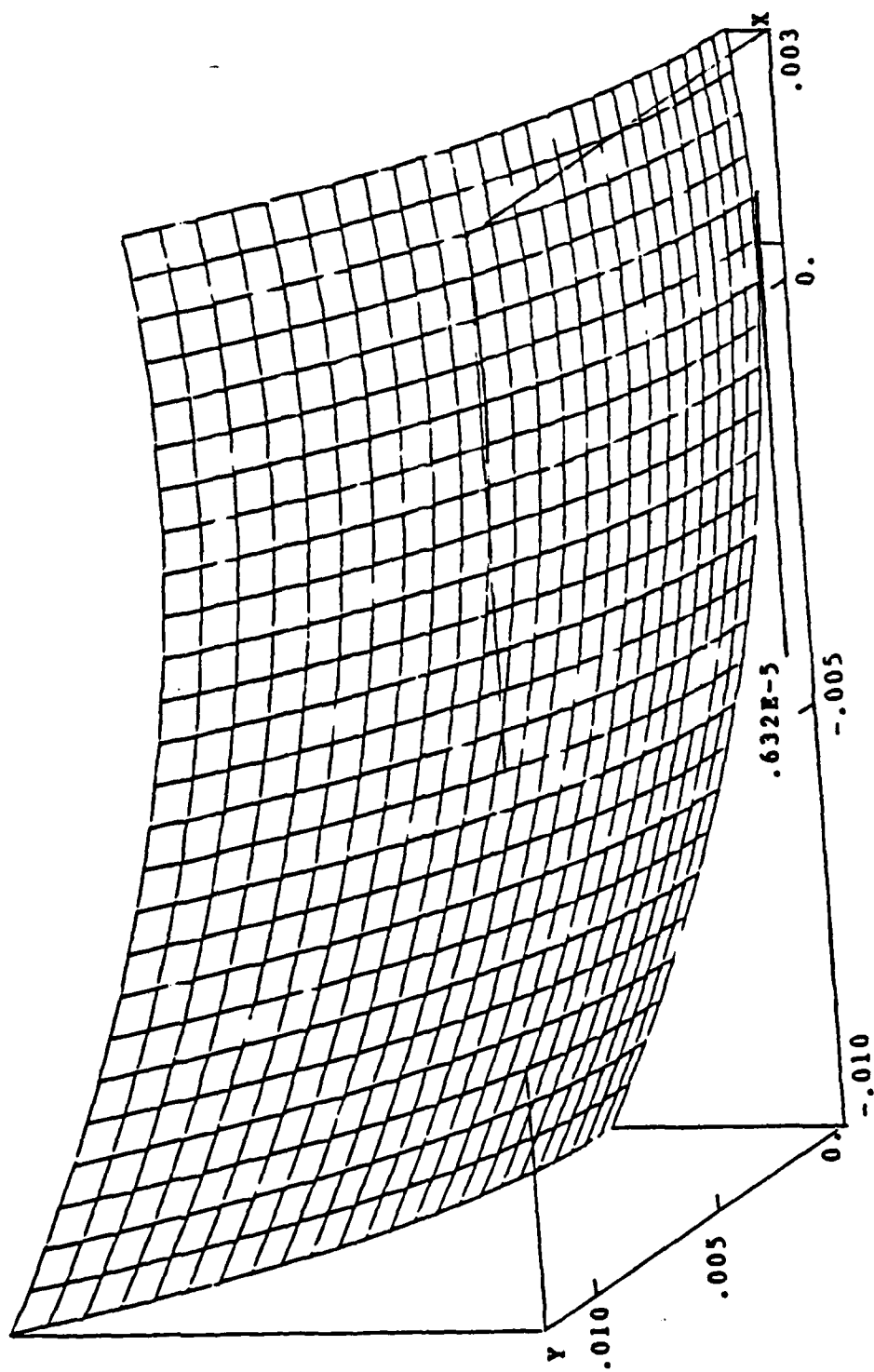


Fig. 5.14 Film Shape Near Contact Center (2-D Biquadratic Elements)

6. ERROR ESTIMATE

Accuracy and error estimation is an important issue in determining the utility and acceptability of any numerical scheme. Since the early 1970's, considerable effort has been expended in developing theories concerning the numerical accuracy of finite element methods. Particularly for elliptical problems, Oden and Reddy^[60] presented an introduction to the mathematical theory of finite elements and Ciarlet^[14] summarized the important results of that period, including some nonlinear cases, see also Oden and Carey^[56]. Some results of finite element methods for nonlinear problems characterized by strongly monotone operators were also obtained [36,53,58]. Oden and Reddy^[59] made a substantial generalization of error estimation methods to problems in nonlinear elasticity which featured a pseudomonotone operator. While a great number of articles have been devoted to fluid dynamics, plates and shells, nonlinear elasticity, etc., error estimates for finite element approximations of nonlinear lubrication problems do not appear to be available.

Li and Dai^[49] made a significant analysis for gas thrust bearings, although elasticity of the bearing was not taken into account. Few other works on the mathematical properties of numerical methods for lubrication problems can be found, either on finite element methods or on finite difference methods.

Here error estimates are derived for approximations of the Reynolds-Hertz equations subject to some additional simplifying assumptions. We first demonstrate the convergence behavior of the approximate solutions in finite dimensional spaces. This provides a framework for the accuracy analysis and then the derivation of an error estimate.

6.1 APPROXIMATIONS OF PENALTY SOLUTIONS IN FINITE-DIMENSIONAL SPACES.

For approximating the functions in Sobolev space V , we introduce a family of finite-dimensional subspaces $\{V_n\}$:

$$V_n \subset V_{n+1} \subset \dots \subset V$$

$$\bigcup_{n \geq 1} V_n \text{ is everywhere dense in } V \quad (6.1)$$

For example, the spaces of interpolation functions in finite element methods can assume such a structure, if we employ a regular refinement of the mesh. Under such approximation, we have

$$\forall v \in V, \exists \text{ a sequence } v_n \in V_n \text{ and } v_n \rightarrow v \quad (6.2)$$

We seek solutions of the approximate penalized problems in the subspaces

V_n :

(P_ϵ^n) : Find $p_\epsilon^n \in V_n$ such that

$$\langle A(p_\epsilon^n), q^n \rangle + \langle (p_\epsilon^n)^-, q^n \rangle / \epsilon = \langle f, q^n \rangle \quad \forall q^n \in V_n \quad (6.3)$$

We denote for the penalized operator,

$$A_\epsilon = A + \Phi, \quad A_\epsilon(q) = A(q) + q^- / \epsilon, \quad \forall q \in H_0^1(\Omega) \quad (6.4)$$

All the reasoning in Chapter 4 applies to the subspaces V_n , so we have

THEOREM 6.1. Operator A_ϵ is bounded, hemicontinuous, coercive, pseudomonotone, and continuous on the subspaces V_n ($n=1, 2, \dots$). And there exist solutions $p_\epsilon^n \in V_n \subset H_0^1(\Omega)$ of (6.3). #

Since $V_n \in V, \forall q^n \in V_n, \langle A_\epsilon(p_\epsilon^n), q^n \rangle = \langle f, q^n \rangle$, and also we have $\langle A_\epsilon(p_\epsilon^n), q^n \rangle = \langle f, q^n \rangle$ because p_ϵ is the solution on V . Thus we obtain the "orthogonality" as follows:

LEMMA 6.2. For the approximate solution p_ϵ^n of (6.3) in the approximate space V_n , we have the orthogonality

$$\langle A_\epsilon(p_\epsilon^n) - A_\epsilon(p_\epsilon^n), q^n \rangle = 0, \quad \forall q^n \in V_n \quad (6.5)$$

THEOREM 6.3. For the approximate solutions $\{p_\epsilon^n\}$ in V_n , there exists

a subsequence of $\{p_\epsilon^n\}$, as $n \rightarrow \infty$, which converges weakly to a $p_\epsilon \in V$, a solution of the original penalized problem (3.27) and $A_\epsilon(p_\epsilon^n) \rightarrow f$ weakly in V' . Furthermore, the subsequence, in fact, converges strongly to p_ϵ in V .

PROOF : For the approximate solutions $p_\epsilon^n \in V_n$,

$$\langle f, p_\epsilon^n \rangle = \langle A_\epsilon(p_\epsilon^n), p_\epsilon^n \rangle = \langle (A + \Phi)(p_\epsilon^n), p_\epsilon^n \rangle, \quad n=1, 2, \dots$$

Coercivity of A_ϵ results in the boundedness of $\{p_\epsilon^n\} \in V = H^1_0$, because inequality (4.37) is valid for A_ϵ in place of A , in view of (4.36). Hence we can find a subsequence, still denoted by $\{p_\epsilon^n\}$, which converges weakly to some p_ϵ in V . Meanwhile $\{A_\epsilon(p_\epsilon^n)\}$ is also bounded, in $V' = H^{-1}$ because of the boundedness of A_ϵ . By the same reasoning, there exists a subsequence of $\{A_\epsilon^n(p_\epsilon^n)\}$, with the same notation, which converges weakly to some g in V' .

We are to prove that p_ϵ is a solution and $g = f$.

For any $q \in V$, we may find a sequence $q^n \in V_n$ such that $q^n \rightarrow q$ in V . p_ϵ^n is the solution in V_n , so

$$\langle A_\epsilon(p_\epsilon^n), q^n \rangle = \langle f, q^n \rangle \rightarrow \langle f, q \rangle$$

Using the boundedness of A_ϵ , we have

$$|\langle A_\epsilon(p_\epsilon^n), q - q^n \rangle| \leq \|A_\epsilon(p_\epsilon^n)\|_{V'} \|q - q^n\|_V \rightarrow 0$$

$$\text{Thus } \langle A_\epsilon(p_\epsilon^n), q \rangle = \langle A_\epsilon(p_\epsilon^n), q - q^n \rangle + \langle A_\epsilon(p_\epsilon^n), q^n \rangle \rightarrow \langle f, q \rangle.$$

On the other hand, the fact that $A_\epsilon(p_\epsilon^n) \rightarrow g$ weakly leads to $\langle A_\epsilon(p_\epsilon^n), q \rangle \rightarrow \langle g, q \rangle$.

This yields $\langle g, q \rangle = \langle f, q \rangle, \quad \forall q \in V$. So $g=f$.

Therefore a similar argument results in

$$\langle A_\varepsilon(p_\varepsilon^n), p_\varepsilon^n - p_\varepsilon \rangle = \langle f, p_\varepsilon^n \rangle - \langle A_\varepsilon(p_\varepsilon^n), p_\varepsilon \rangle \rightarrow \langle f, p_\varepsilon \rangle - \langle g, p_\varepsilon \rangle = 0.$$

This means that the property P (3.15) is satisfied by A_ε with the sequence $\{p_\varepsilon^n\}$. A_ε being pseudomonotone gives

$$\lim_{n \rightarrow \infty} \langle A_\varepsilon(p_\varepsilon^n), p_\varepsilon^n - q \rangle \geq \langle A_\varepsilon(p_\varepsilon), p_\varepsilon - q \rangle, \quad \forall q \in V.$$

$$\text{From } \langle A_\varepsilon(p_\varepsilon^n), p_\varepsilon^n \rangle = \langle f, p_\varepsilon^n \rangle \rightarrow \langle f, p_\varepsilon \rangle, \langle A_\varepsilon(p_\varepsilon^n), q \rangle \rightarrow \langle g, q \rangle = \langle f, q \rangle$$

we have $\langle f, p_\varepsilon - q \rangle \geq \langle A_\varepsilon(p_\varepsilon), p_\varepsilon - q \rangle, \quad \forall q \in V.$

For any $w \in V$, setting $q = p_\varepsilon - \theta w \Rightarrow \langle f, \theta w \rangle \geq \langle A_\varepsilon(p_\varepsilon), \theta w \rangle$. Equivalently $\langle A_\varepsilon(p_\varepsilon), w \rangle = \langle f, w \rangle$ since θ can be chosen positive or negative. This establishes that p_ε is a solution of (3.27).

Going a step further, if we take a sequence $r^n \in V_n, r^n \rightarrow p_\varepsilon$ in V , then from boundedness, we have

$$|\langle A_\varepsilon(p_\varepsilon) - A_\varepsilon(p_\varepsilon^n), p_\varepsilon - r^n \rangle| \leq \|A_\varepsilon(p_\varepsilon) - A_\varepsilon(p_\varepsilon^n)\|_V \|p_\varepsilon - r^n\|_V \rightarrow 0.$$

By virtue of (4.18) and monotonicity of Φ (4.36), we have

$$\begin{aligned} & \langle A_\varepsilon(p_\varepsilon) - A_\varepsilon(p_\varepsilon^n), p_\varepsilon - p_\varepsilon^n \rangle \\ & \geq C_1 \|p_\varepsilon - p_\varepsilon^n\|_1^2 - C_\eta \|p_\varepsilon - p_\varepsilon^n\|_{L^p} \|p_\varepsilon - p_\varepsilon^n\|_{H^1}. \end{aligned}$$

Orthogonality (6.5) gives

$$\langle A_\varepsilon(p_\varepsilon) - A_\varepsilon(p_\varepsilon^n), p_\varepsilon - r^n \rangle = \langle A_\varepsilon(p_\varepsilon) - A_\varepsilon(p_\varepsilon^n), p_\varepsilon - p_\varepsilon^n \rangle.$$

Again the compact embedding of H^1 in L^q results in $p_\varepsilon^n \rightarrow p_\varepsilon$ strongly in L^q .

Finally we obtain $\|p_\varepsilon - p_\varepsilon^n\|_{H^1} \rightarrow 0.$

#

6.2 APPROXIMATION OF VARIATIONAL INEQUALITIES IN FINITE-DIMENSIONAL SPACES.

We now return to the variational inequality (3.27) and consider a finite-dimensional approximation of problem (P).

As before, let $V_n \subset V_{n+1} \subset \dots \subset V$ and $\cup V_n$ be everywhere dense in V . Denote

$$K_n = K \cap V_n = \{q \in V_n \mid q \geq 0 \text{ a.e. in } \Omega\}$$

$$K_n \subset K_{n+1} \subset \dots \subset K, \quad \overline{\cup K_n} = K \quad (6.6)$$

We also have the approximation property:

$$\forall q \in K, \exists \text{ a sequence } q^n \in K_n, \quad q^n \rightarrow q \text{ in } V \text{ norm} \quad (6.7)$$

In fact, when we consider finite element approximations, V_n are the spaces of piecewise polynomials. Generally, for $q \in K, q \geq 0$ a.e. in Ω , we can get a sequence $q^n \in V_n, q^n \rightarrow q$ in the V norm; q^n may not be in K_n , however $\Omega^n_0 = \{x \in \Omega \mid q^n \leq 0\}$ must shrink to a set of zero measure. So we may use $(q^n)^+ \in K_n$ to approximate q .

The K_n are closed and convex subsets of V_n . Our approximation is characterized as the problem

(P^n) : Find $p^n \in K_n$, such that

$$\langle A(p^n) - f, q^n - p^n \rangle \geq 0 \quad \forall q^n \in K_n \quad (6.8)$$

Making the same argument as in proving Theorem 4.13 on V_n , we have directly the convergence property.

THEOREM 6.4. In the approximate spaces V_n , when $\varepsilon \rightarrow 0$, there exists a subsequence of $\{p_\varepsilon^n\}$, the solutions of the penalized problem (6.3), which

converges to some $p^n \in K_n$ in V , the solution of (6.8). #

Recall that in finite-dimensional spaces, weak convergence is equivalent to strong convergence.

On the other hand we have for the approximation for (3.27):

THEOREM 6.5. There exists a subsequence of $\{p^n\}$, the approximate solutions of variational inequality in finite-dimensional spaces (6.8), which converges to some $p \in K$ in the V norm as $n \rightarrow \infty$. Moreover, p is a solution of problem (P), the variational inequality (3.27).

PROOF : We may take p^n as the limit of a convergent sequence $\{p_\varepsilon^n\}$ of penalty solutions in finite-dimensional spaces. The boundedness of $\{p_\varepsilon^n\}$, by virtue of (4.38):

$$\|p_\varepsilon^n\|_{H^1} \leq (C_2 + \|f\|_V)/C_1$$

and convergence of $p_\varepsilon^n \rightarrow p^n$ as $\varepsilon \rightarrow 0$, lead to the boundedness of $\{p^n\}$.

Thus we can find a subsequence, still denote $\{p^n\}$, which converges weakly to some $p \in V$.

The space $V = H_0^1 \subset L^2$ compactly, so $p^n \rightarrow p$ strongly in L^2 , $\Rightarrow p^n \rightarrow p$ a.e. in Ω ; $p^n \in K_n$, $p^n \geq 0$ a.e. in $\Omega \Rightarrow p \geq 0$ a.e. in Ω . Hence $p \in K$.

Let us choose an approximate sequence q^n for p in K_n , $q^n \rightarrow p$ in V . Then we have

$$\begin{aligned} \langle A(p^n), p^n - p \rangle &= \langle A(p^n), p^n - q^n \rangle + \langle A(p^n), q^n - p \rangle \\ &\leq \langle f, p^n - q^n \rangle + \langle A(p^n), q^n - p^n \rangle \\ &= \langle f, p^n - p \rangle + \langle A(p^n) - f, q^n - p \rangle \end{aligned}$$

$p^n \rightarrow p$ weakly, but $q^n \rightarrow p$ strongly and $\{A(p^n)\}$ is bounded. Therefore

$$\overline{\lim} \langle A(p^n), p^n - p \rangle \leq 0 \quad (6.9)$$

That means A and $\{p^n\}$ satisfy property P (3.15). A is pseudomonotone and thus we have

$$\liminf \langle A(p^n), p^n - q \rangle \geq \langle A(p), p - q \rangle, \quad \forall q \in V.$$

Then, for any $q \in K$, we may construct as an approximation, $q^n \in K_n$, $q^n \rightarrow q$ in V ,

$$\begin{aligned} \langle A(p), p - q \rangle &\leq \liminf \langle A(p^n), p^n - q \rangle \\ &= \liminf [\langle A(p^n), p^n - q^n \rangle + \langle A(p^n), q^n - q \rangle] \leq \liminf \langle f, p^n - q^n \rangle = \langle f, p - q \rangle \end{aligned}$$

i.e., $\langle A(p) - f, q - p \rangle \geq 0, \quad \forall q \in K.$

Finally we can prove that the convergence is in the H^1 norm.

Lemma 4.7 gives

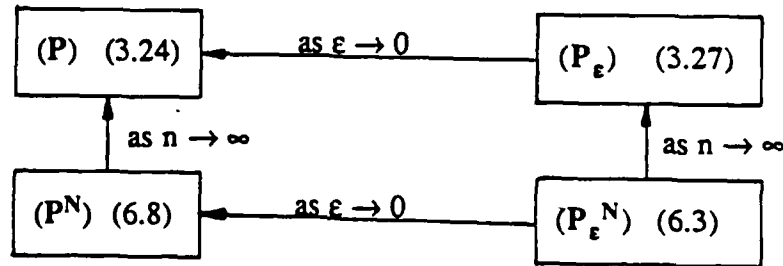
$$\langle A(p^n) - A(p), p^n - p \rangle \geq C_1 \|p^n - p\|_1^2 - C_\eta \|p^n - p\|_{L^q} \|p^n - p\|_{H^1}.$$

Recalling (6.9), we have $\overline{\lim} \langle A(p^n), p^n - p \rangle \leq 0$. So

$$\begin{aligned} C_1 (\|p^n - p\|_{H^1})^2 &\leq C_\eta \|p^n - p\|_{L^q} \|p^n - p\|_{H^1} + \overline{\lim} \langle A(p^n) - A(p), p^n - p \rangle \\ &\leq C_\eta 2\eta \|p^n - p\|_{L^q} + \overline{\lim} \langle -A(p), p^n - p \rangle \rightarrow 0 \end{aligned}$$

Note that we have used the boundedness of $\{p^n\}$. So we can conclude that $p^n \rightarrow p$ strongly in H^1 . #

Up to now we have constructed the finite-dimensional approximations for both the variational inequality and the penalized problem. Also we have established the convergence relations (all in the H^1 -norm), summarized in the following diagram.



6.3 AN A-PRIORI ERROR ESTIMATE.

Making use of inequality (4.18) in Lemma 4.7 again, along with the monotonicity of the operator Φ and the orthogonality for the approximate solution p_ε^n , we have

$$\begin{aligned} & C_1(\|p_\varepsilon - p_\varepsilon^n\|_{H^1})^2 - C_\eta \|p_\varepsilon - p_\varepsilon^n\|_{L^q} \|p_\varepsilon - p_\varepsilon^n\|_{H^1} \\ & \leq \langle A(p_\varepsilon) - A(p_\varepsilon^n), p_\varepsilon - p_\varepsilon^n \rangle \leq \langle A_\varepsilon(p_\varepsilon) - A_\varepsilon(p_\varepsilon^n), p_\varepsilon - p_\varepsilon^n \rangle \\ & = \langle A_\varepsilon(p_\varepsilon) - A_\varepsilon(p_\varepsilon^n), p_\varepsilon - q^n \rangle, \quad \forall q^n \in V_n \end{aligned}$$

By the continuity of A in Theorem 4.9,

$$\|A(p_\varepsilon) - A(p_\varepsilon^n)\|_{L^1} \leq C_A \|p_\varepsilon - p_\varepsilon^n\|_{H^1}.$$

Note that, $\forall q_1, q_2 \in V$

$$\|\Phi(q_1) - \Phi(q_2)\|_{L^1} = \sup_{\|r\|_1=1} \langle q_1^- - q_2^-, r \rangle / \varepsilon \|r\|_{H^1}$$

Recalling, the process in proving Lemma 4.11, and $\Omega_\alpha = \{x \in \Omega \mid q_\alpha \geq 0\}$,

$\Omega_\alpha' = \{x \in \Omega \mid q_\alpha < 0\}$, we have

$$\begin{aligned} & (\|q_1^- - q_2^-\|_{L^2})^2 = \int_\Omega (q_1^- - q_2^-)^2 d\Omega \\ & = \int_{\Omega_1' \cap \Omega_2'} (q_1 - q_2)^2 d\Omega + \int_{\Omega_1' \cap \Omega_2} q_1^2 d\Omega + \int_{\Omega_1 \cap \Omega_2'} q_2^2 d\Omega \end{aligned}$$

in $\Omega_1' \cap \Omega_2'$, $q_1 < 0$ and $q_2 \geq 0$, $|q_1| \leq |q_1 - q_2|$; in $\Omega_1' \cap \Omega_2$, $|q_2| \leq |q_1 - q_2|$. Thus

$$\|q_1^- - q_2^-\|_0^2 \leq \int_{\Omega_1' \cap \Omega_2'} + \int_{\Omega_1' \cap \Omega_2} + \int_{\Omega_1 \cap \Omega_2'} (q_1 - q_2)^2 d\Omega \leq \|q_1 - q_2\|_0^2.$$

Therefore $\|\Phi(q_1) - \Phi(q_2)\|_{L^1} \leq \|q_1 - q_2\|_0 / \varepsilon \leq \|q_1 - q_2\|_1 / \varepsilon$. Then we arrive at an inequality.

LEMMA 6.4. For the approximate solutions $\{p_\varepsilon^n\}$ of problem (P_ε^n)

(6.3) for $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$, such that

$$C_1 \|p_\varepsilon - p_\varepsilon^n\|_{H^1} - C_\eta \|p_\varepsilon - p_\varepsilon^n\|_{L^q} \leq C_\varepsilon \|p_\varepsilon - r^n\|_{H^1}, \quad \forall r^n \in V_n \quad (6.6)$$

$$C_\varepsilon \leq C_A + 1/\varepsilon \quad (6.7)$$

where C_1 , C_η and C_A are defined in (4.16), (4.21) and (4.28) respectively. #

Now we use parameter h traditionally to indicate the mesh size in the finite element method, and we construct a family of finite element approximation spaces with mesh parameter h_n as follows:

$$V_{h_1} \subset V_{h_2} \subset \dots \subset V, \quad 1 > h_1 > h_2 > \dots > 0$$

$$\cup V_{h_n} \text{ is everywhere dense in } V \text{ with } h_n \rightarrow 0 \quad (6.8)$$

According to finite element interpolation theory (see, [14] or [56]), if complete polynomials of degree not larger than k are contained in $H^r(\Omega)$, generally we have a constant $C_1(m, k, \Omega) > 0$ such that $\forall v \in H^{k+1}(\Omega)$,

$$\|v - \Pi_h v\|_{H^m} \leq C_1 h^{k+1-m} |v|_{H^{k+1}(\Omega)}, \quad 0 \leq m \leq \min(1, r)$$

$$\left[\sum_{e=1}^E (\|v - \Pi_h v\|_{H^m(\Omega_e)})^2 \right]^{1/2} \leq C_1 h^{k+1-m} |v|_{H^{k+1}(\Omega)}, \quad 2 \leq m \leq \min(k+1, r) \quad (6.9)$$

where $\Pi_h : H^{k+1} \rightarrow S^h \subset H_m$ is the interpolation operator with S^h denoting the finite element space, and $|\cdot|$ denoting the seminorm.

When these results are valid, we are able to establish an a-priori error estimate for the approximate solutions of penalized problem in finite-dimensional spaces V_h . Using (6.6) and, setting $m=1$, we obtain

THEOREM 6.5. If the penalty solution $p_\varepsilon \in H^r(\Omega)$, $r \geq 1$ and the shape functions of finite elements contain P_k , the complete polynomials of degrees $\leq k$,

$k \geq 1$; then under the assumption (4.33), $C_1 - C_\eta C_q > 0$, and for the regular affine family of finite elements, there exists a constant $C(\epsilon, m, k, \Omega, \eta) > 0$, such that for the approximate solutions p_ϵ^h on V_h , we have the error estimate

$$\begin{aligned} \|p_\epsilon - p_\epsilon^h\|_{H^1} &\leq Ch^\mu |p_\epsilon|_{H^{k+1}} \\ \mu &= \min(k, r-1) \end{aligned} \quad (6.10) \quad \#$$

COROLLARY 6.6. If the penalty solution p_ϵ is smooth enough to be an element of H^r , $r \geq k+1$, under the conditions in Theorem 6.5, we have the error estimate

$$\|p_\epsilon - p_\epsilon^h\|_{H^1} \leq Ch^k |p_\epsilon|_{H^{k+1}} \quad (6.11)$$

Particularly, for linear elements, $k=1$, and for quadratic elements $k=2$, (6.11) yields

$$\|p_\epsilon - p_\epsilon^h\|_{H^1} \leq Ch |p_\epsilon|_{H^2} \quad (\text{linear elements}) \quad (6.12)$$

$$\|p_\epsilon - p_\epsilon^h\|_{H^1} \leq Ch^2 |p_\epsilon|_{H^3} \quad (\text{quadratic elements}) \quad (6.13) \quad \#$$

REMARK 1. For one-dimensional problems, Theorem 4.16 and 4.17 demonstrate that $p_\epsilon \in H^2(\Omega)$, so at least (6.12) is realistic.

REMARK 2. For one-dimensional penalized problems, it is proved that $p_\epsilon \in C^1[a, b]$. By (4.47) and (4.50), we may obtain

$$h(p_\epsilon) - h(p_\epsilon^h) = \frac{2}{\pi E'} \int_a^b \ln \left(\frac{\xi - x_0}{\xi - x} \right)^2 (p_\epsilon(\xi) - p_\epsilon^h(\xi)) d\xi$$

$$h'(p_\epsilon) - h'(p_\epsilon^h) = \frac{2}{\pi E'} \int_a^b -\ln(\xi - x)^2 (p_\epsilon'(\xi) - (p_\epsilon^h)'(\xi)) d\xi$$

Thus we have

$$\|h_\epsilon - h_\epsilon^h\|_{H^s} \leq C \|p_\epsilon - p_\epsilon^h\|_{H^s}, \quad s = 0, 1 \quad (6.14)$$

That means the film thickness has at least the same rate of convergence as the pressure.

Moreover, for the load $w = \int p d\Omega$, we have

$$|w_\epsilon - w_\epsilon^h| = |\int p_\epsilon d\Omega - \int p_\epsilon^h d\Omega| \leq C \|p_\epsilon - p_\epsilon^h\|_{L^2}. \quad (6.15)$$

Thus the convergence of load is as at least fast as the pressure measured in the L^2 -norm.

6.4 NUMERICAL EXPERIMENTS ON ERROR ESTIMATES FOR FINITE ELEMENT SOLUTIONS.

A series of tests on finite element solutions have been performed. For simplicity, uniform meshes are employed. A very fine mesh is taken for obtaining a solution p_ϵ^* in place of the exact solution. And then the analysis is based on the comparisons between this solution and the solutions obtained from the coarse meshes.

6.4.1. LINE CONTACT PROBLEMS.

Figure 6.1 shows the computed error behavior for a light load case of line contact, computed with the linear elements. It is evident that the logarithm relation between $\|p_\epsilon^* - p_\epsilon^h\|_{H^1}$ and h , the mesh parameter, is linear with slope 1. That is, a first order convergence is obtained, and the prediction (6.12) (see Corollary 6.6 and Remark 1) is verified. In fact, shown in Figs. 6.2 to 6.4, the result holds for a quite wide range of loads.

Figures 6.5 to 6.8 show the analysis with quadratic elements. For a certain range of loads, the rates of convergence for pressure are $O(h^2)$, as predicted in (6.13).

For classical linear elliptic problems, such results are optimal. Since the operator of the governing equation is only pseudomonotone, the convergent rates are almost optimal only for a certain range of loads.

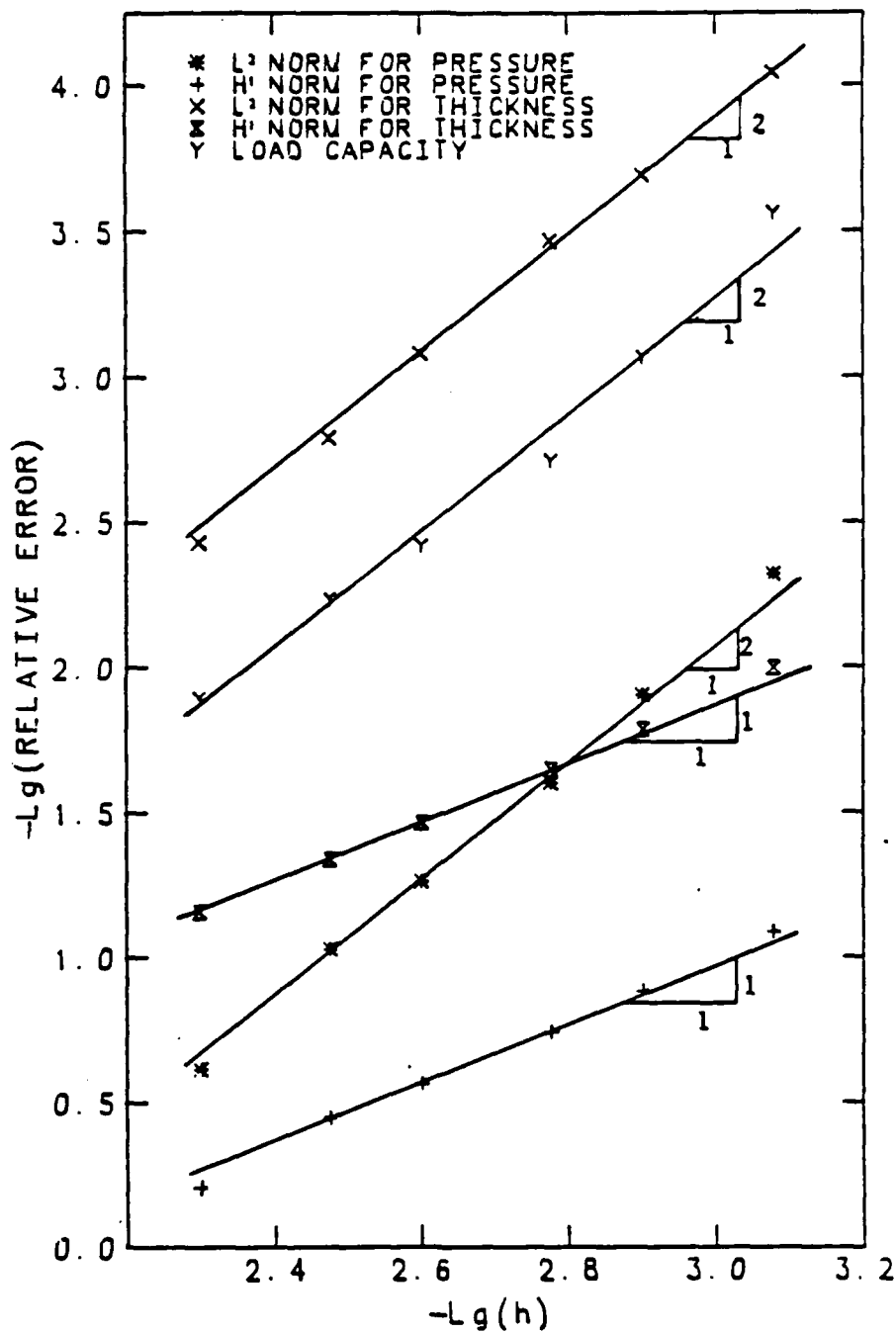


Fig. 6.1 Error Analysis (1-D Linear Elements, $H_0=0.2E-4$)
Line Contact ($W=0.166E-5$)

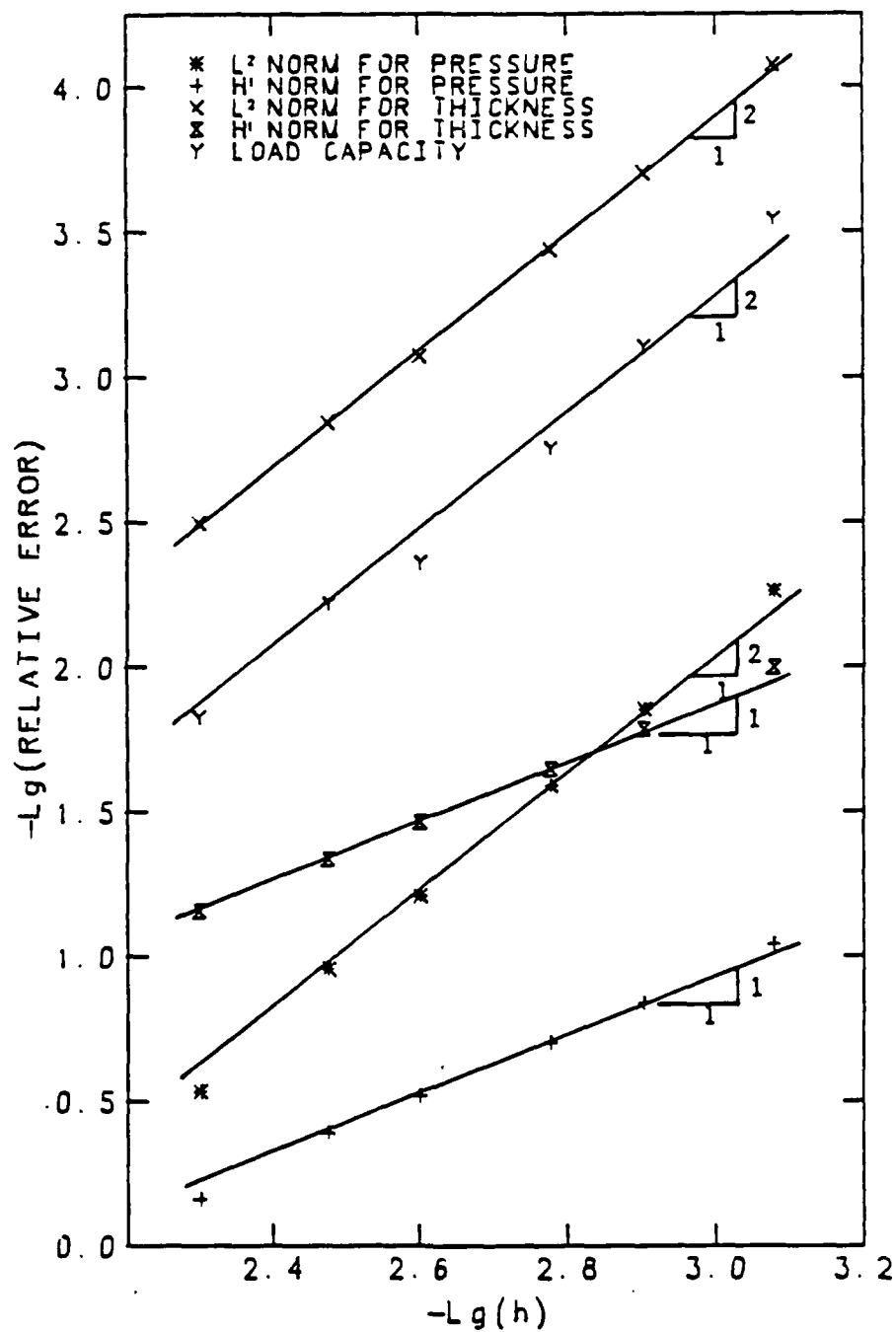


Fig. 6.2 Error Analysis (1-D Linear Elements, $H_0=0.$)
Line Contact ($W=0.304E-5$)

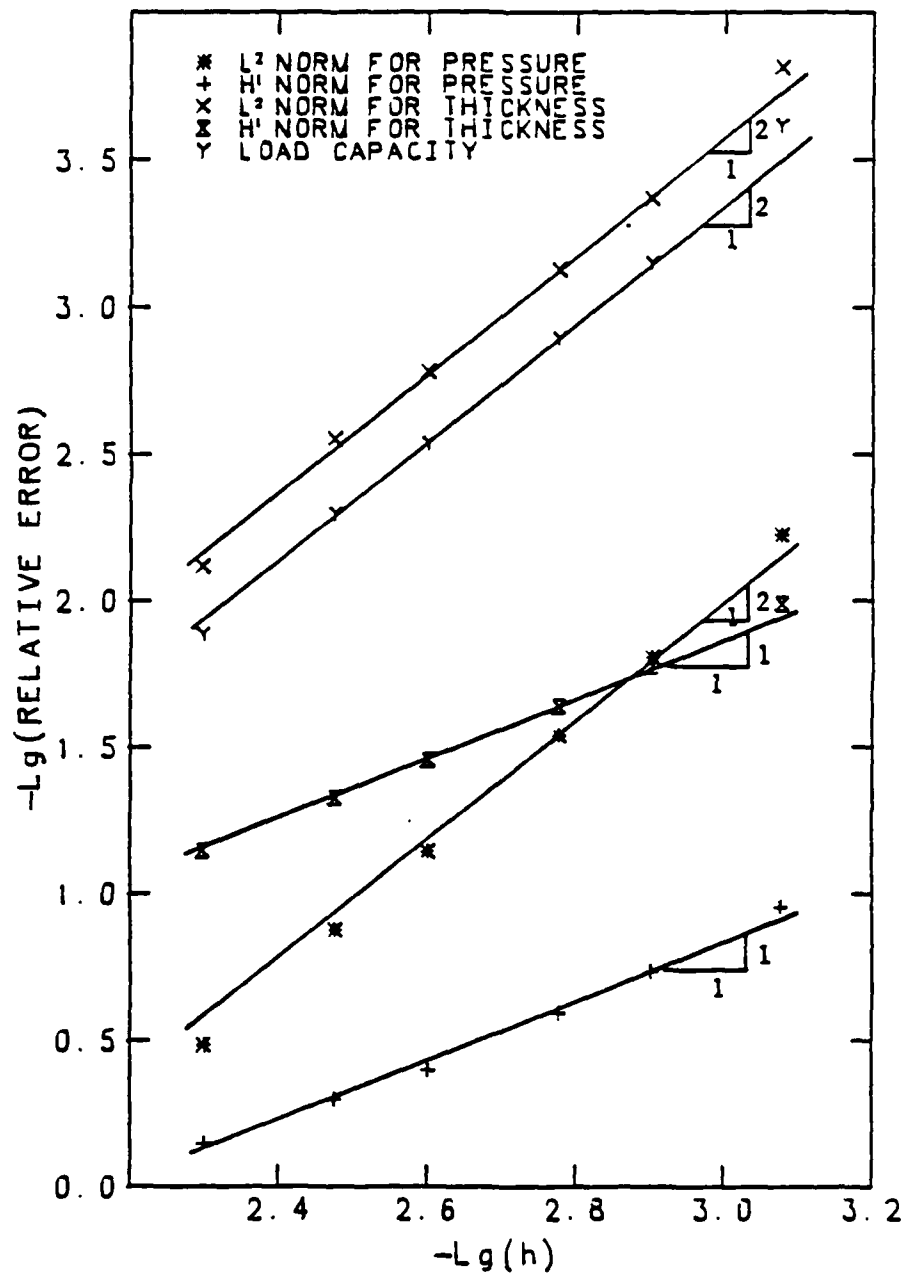


Fig. 6.3 Error Analysis (1-D Linear Elements, $H_0 = -0.2E-4$)
Line Contact ($W = 0.515E-5$)

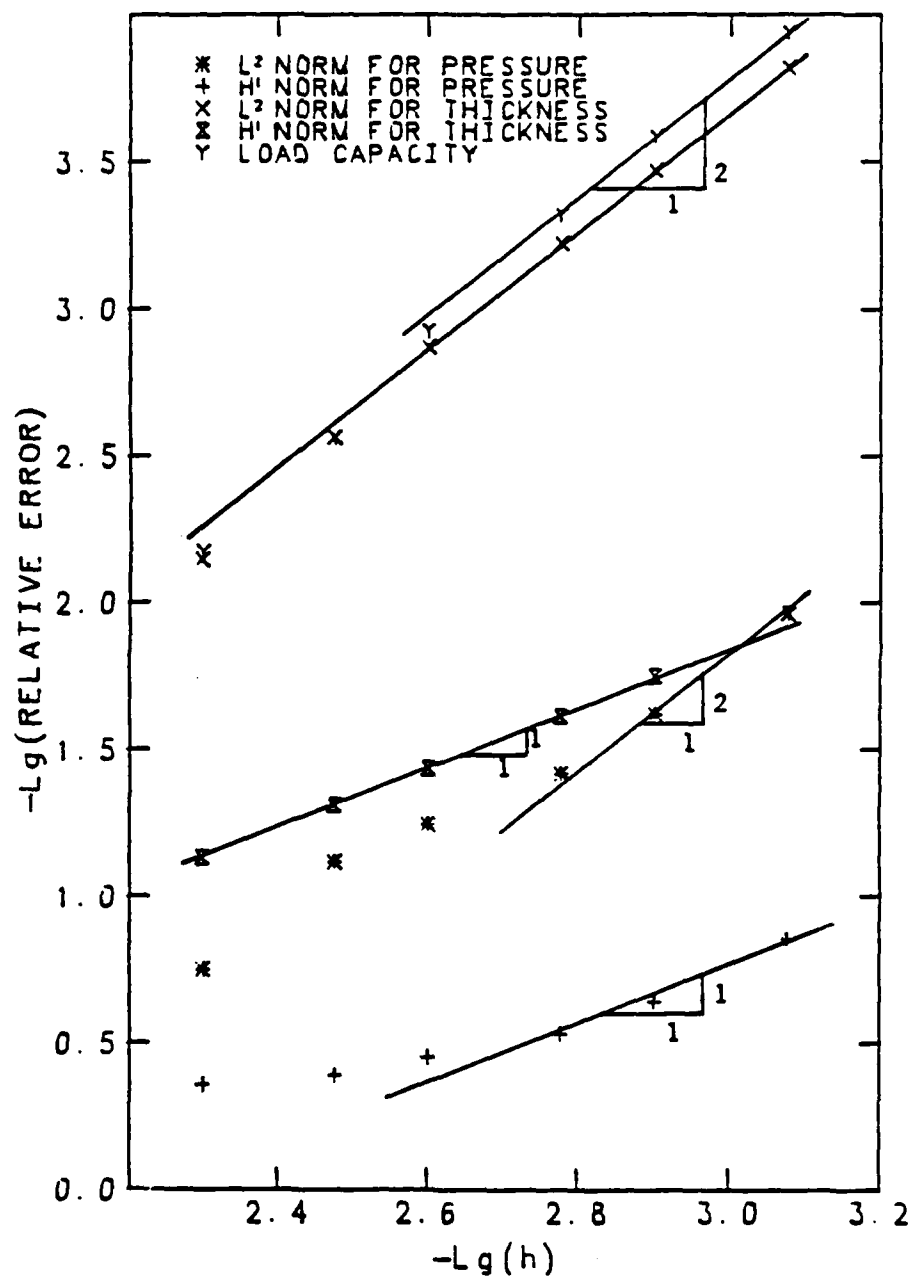


Fig. 6.4 Error Analysis (1-D Linear Elements, $H_0 = -0.5E-4$)
Line Contact ($W = 0.888E-5$)

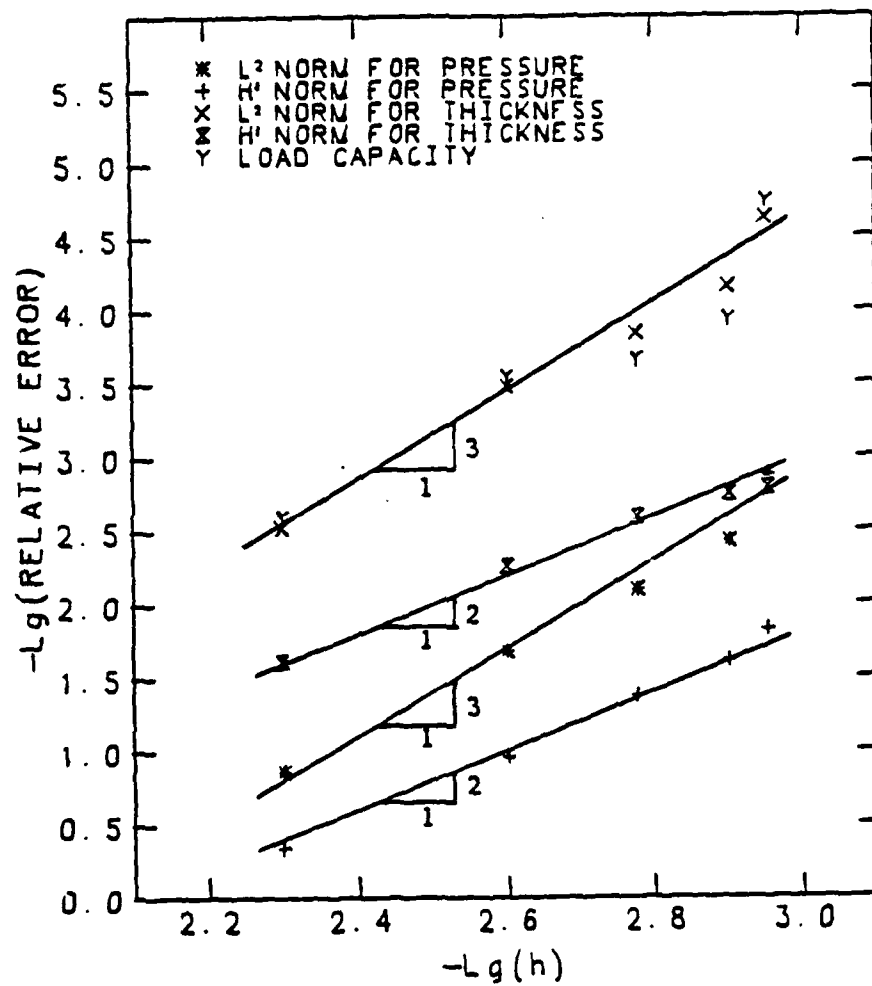


Fig. 6.5 Error Analysis (1-D Quadratic Elements, $H_0=0.$)
Line Contact ($W=0.305E-5$)

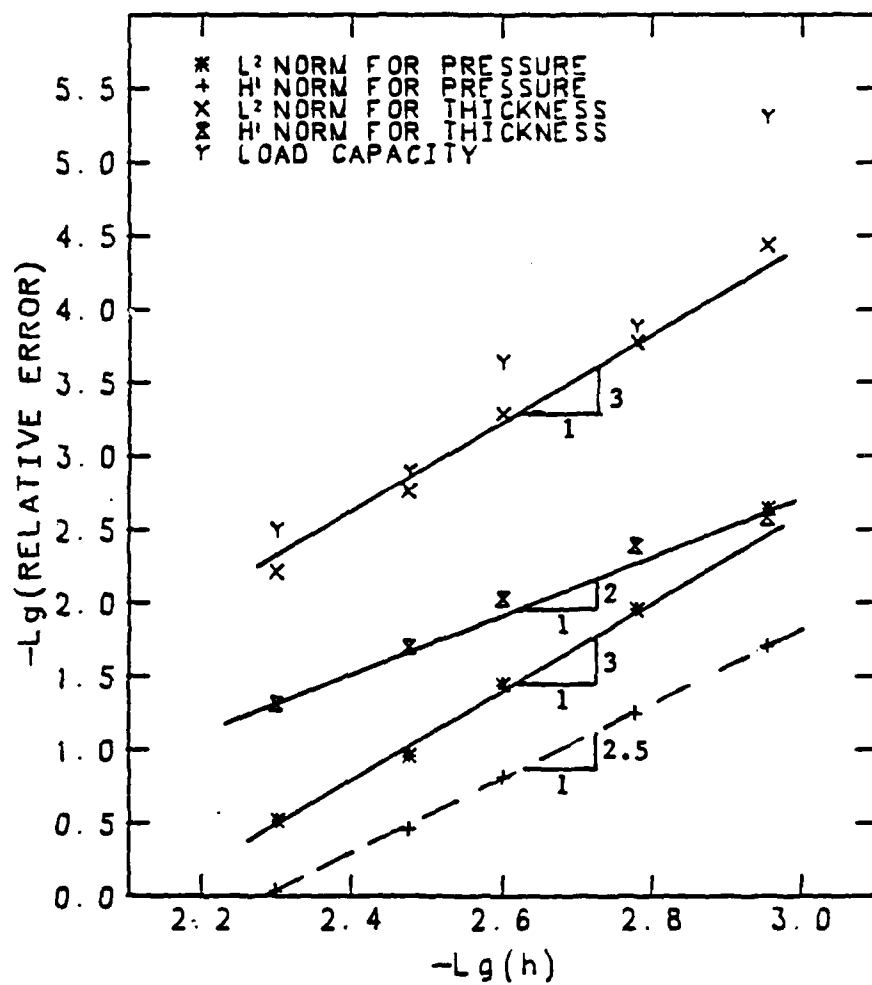


Fig. 6.6 Error Analysis (1-D Quadratic Elements, $H_0 = -0.1E-4$)
Line Contact ($W = 0.404E-5$)

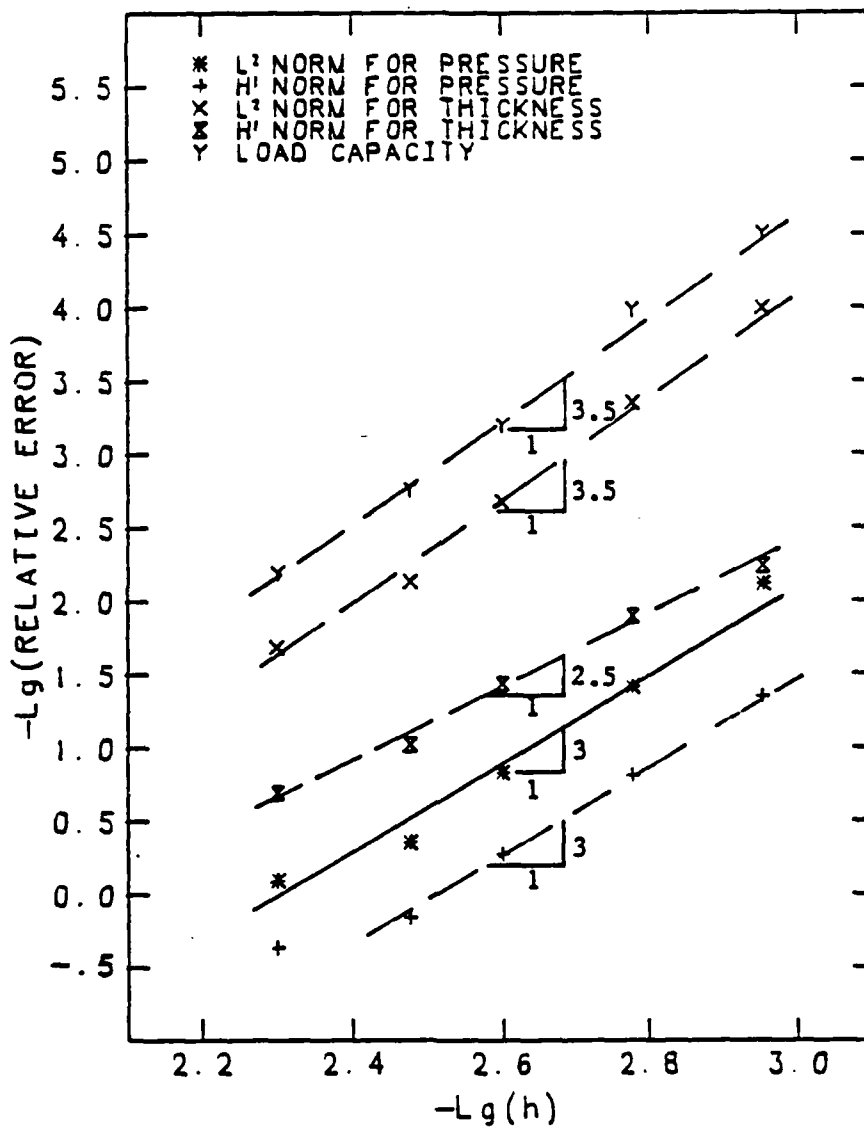


Fig. 6.7 Error Analysis (1-D Quadratic Elements, $H_0 = -0.4E-4$)
Line Contact ($W = 0.762E-5$)

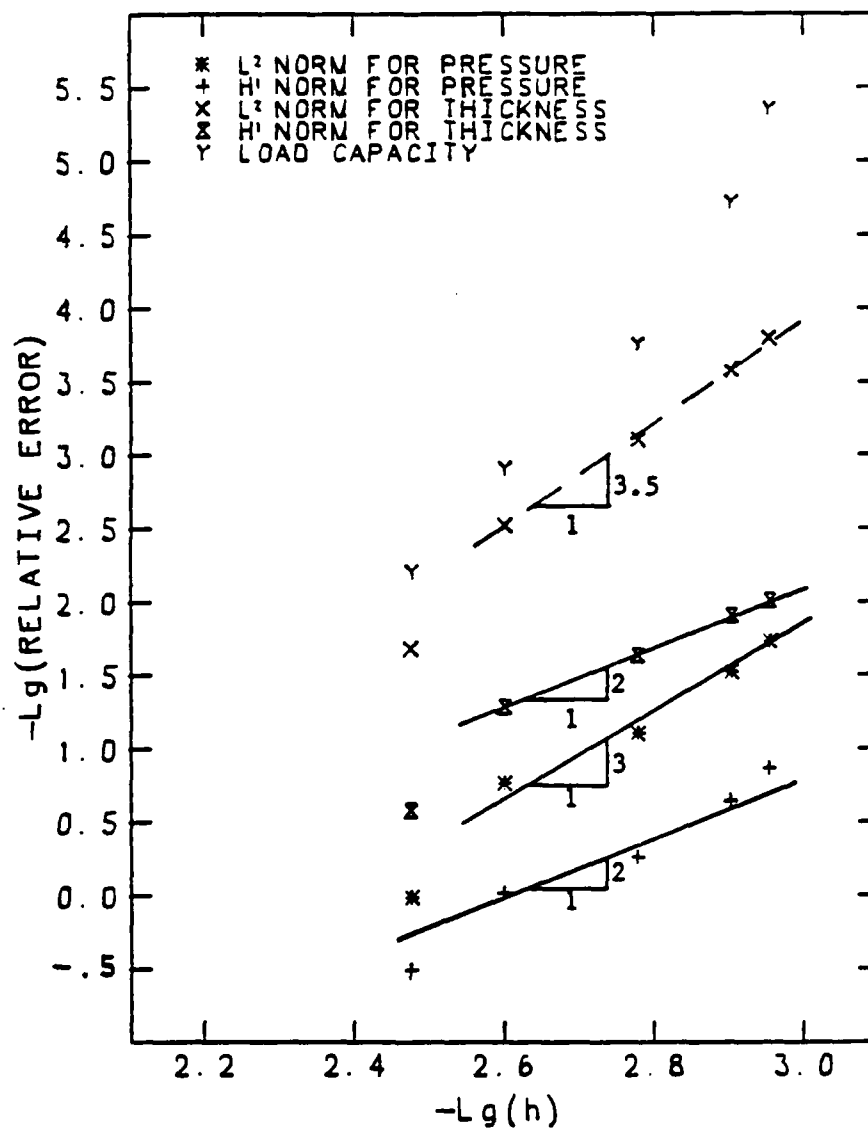


Fig. 6.8 Error Analysis (1-D Quadratic Elements, $H_0 = -0.6E-4$)
Line Contact ($W = 101E-4$)

6.4.2. POINT CONTACT PROBLEMS.

For two-dimensional point contact problems, Figs. 6.9 to 6.11 show that (6.12) also holds for certain cases with bilinear elements, but a deterioration in the rate of convergence is observed for slightly heavier loads as shown in Fig. 6.12. A similar situation appears for the solutions computed with biquadratic elements, shown in Figs. 6.13 to 6.16. This indicates that condition (4.33) does not hold all the time or the solutions are not smooth enough to be in the Sobolev spaces of higher order.

REMARK. For linear elliptic problems, the Aubin–Nitsche method (see [56]) provides higher order convergence when the error is measured in the Sobolev spaces of lower orders. Here it is interesting to notice that the similar behavior is observed. Figures 6.1 to 6.16 show that in many cases, the pressure obtains second order convergence with linear elements, and third order convergence with quadratic elements, when measured in the L^2 -norm, just one order higher than those in the H^1 -norm, as is the case in linear elliptic problems.

6.5 APPLICATION OF ADAPTIVE METHODS.

The solution of the elastohydrodynamic lubrication problem usually exhibits a large gradient in the pressure distribution. For improving quality of approximations of the pressure, more nodes can be placed in the region where the pressure varies rapidly. Alternately, one can use higher-order elements in that region. However, such regions may vary case by case and the use of very fine uniform meshes is not feasible as it may require excessive computer storage and time, particularly in case of point contact problems. To then obtain good approximations, the use of adaptive finite element methods to automatically produce a suitable refinement is thus an attractive alternative. Such schemes have been applied to many engineering problems in recent years, e.g., see [23,24,61]. The basic ideas and schemes of adaptive method can be found in [4,5,39], and the recent advances and applications are summarized in [6].

Generally, there are three types of adaptive methods: h -methods, p -methods and moving mesh methods. Here we shall apply h -methods, which involve adaptive mesh refinement, to the elastohydrodynamic lubrication problems.

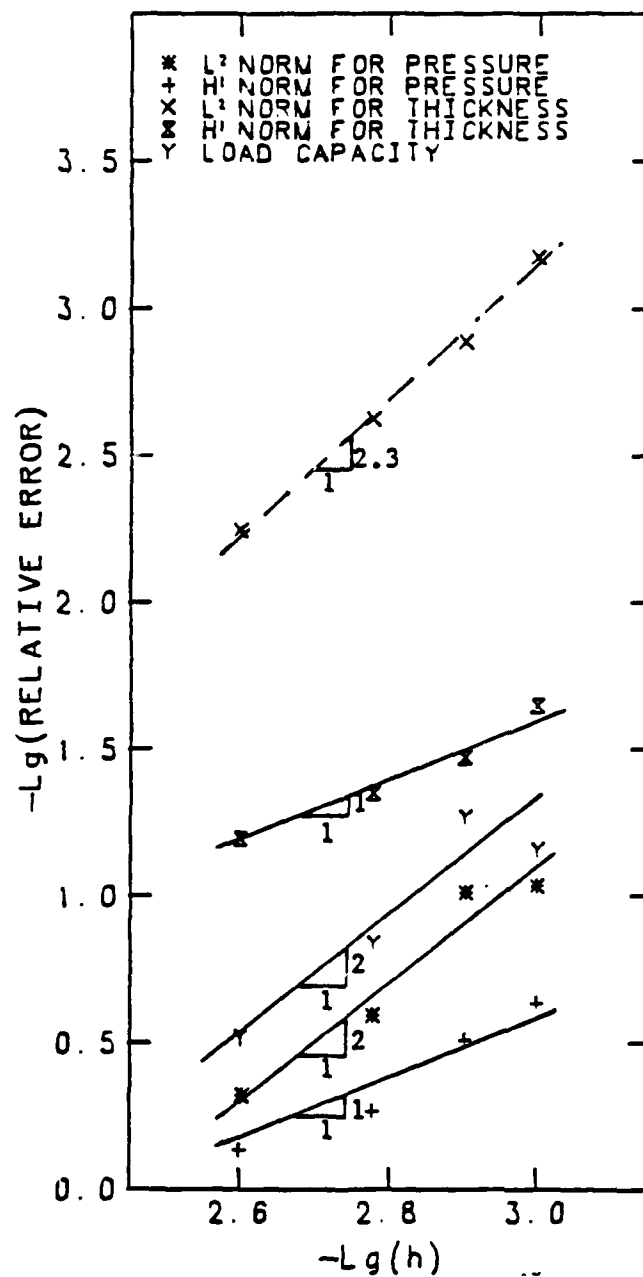


Fig. 6.9 Error Analysis (2-D Bilinear Elements, $H_0=0.5E-5$)
Point Contact ($W=0.680E-8$)

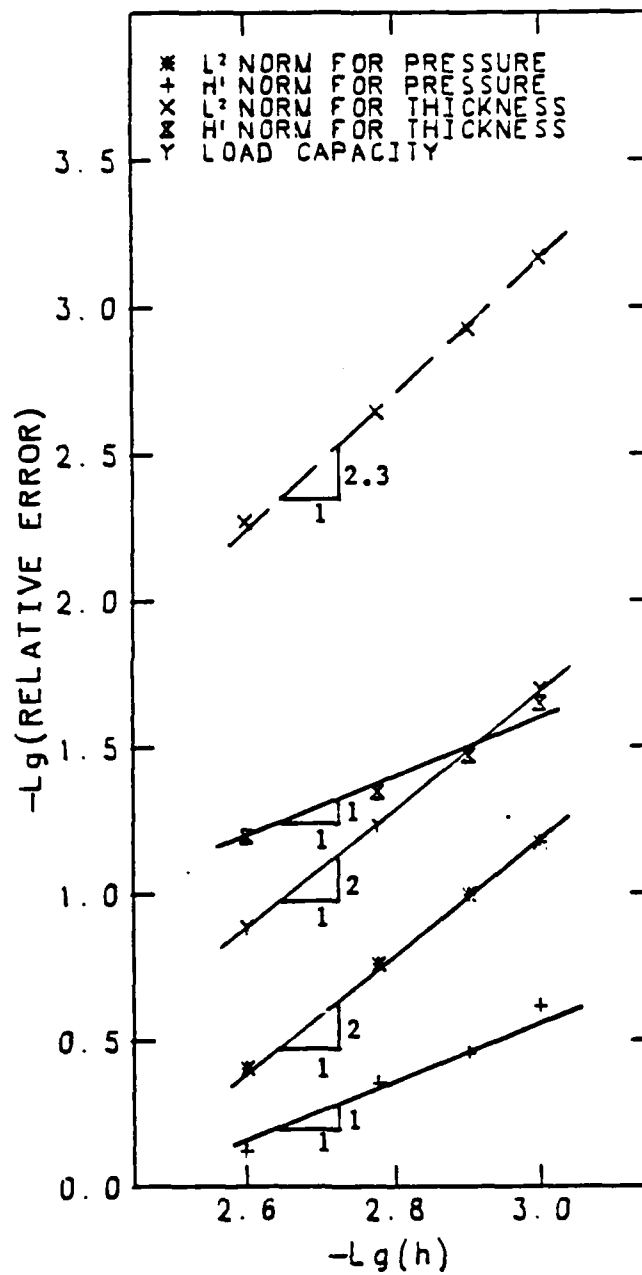


Fig. 6.10 Error Analysis (2-D Bilinear Elements, $H_0=0.3E-5$)
Point Contact ($W=0.788E-8$)

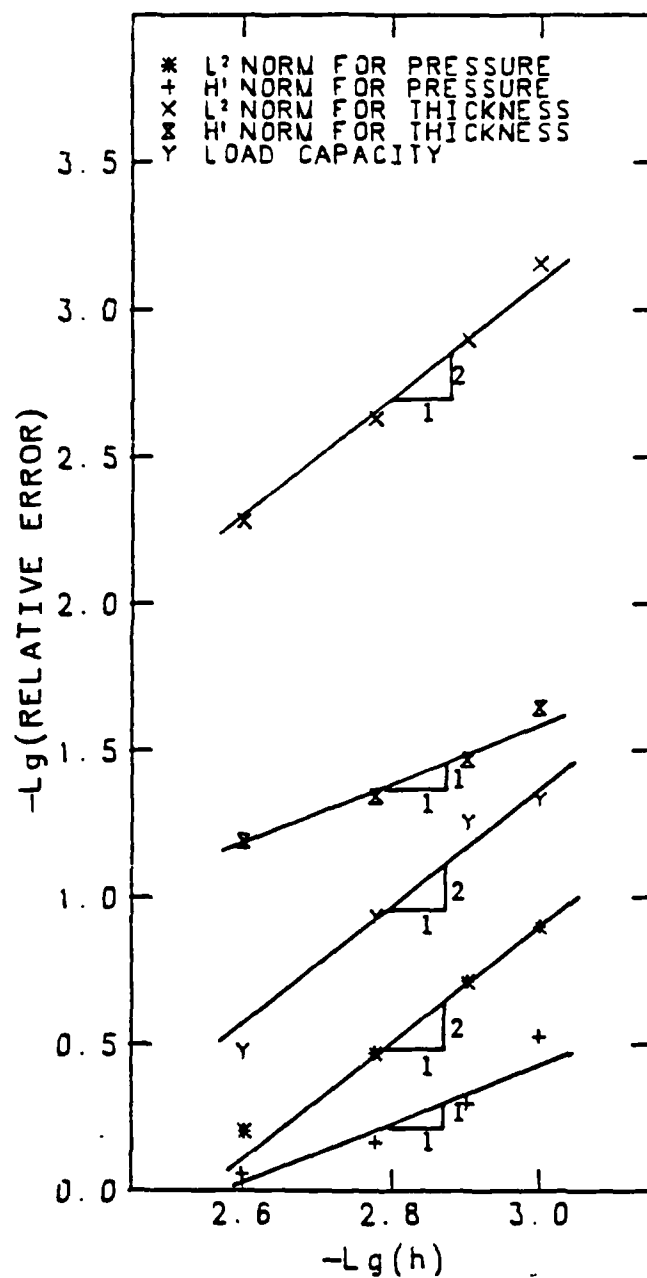


Fig. 6.11 Error Analysis (2-D Bilinear Elements, $H_0=0.1E-5$)
Point Contact ($W=0.126E-7$)

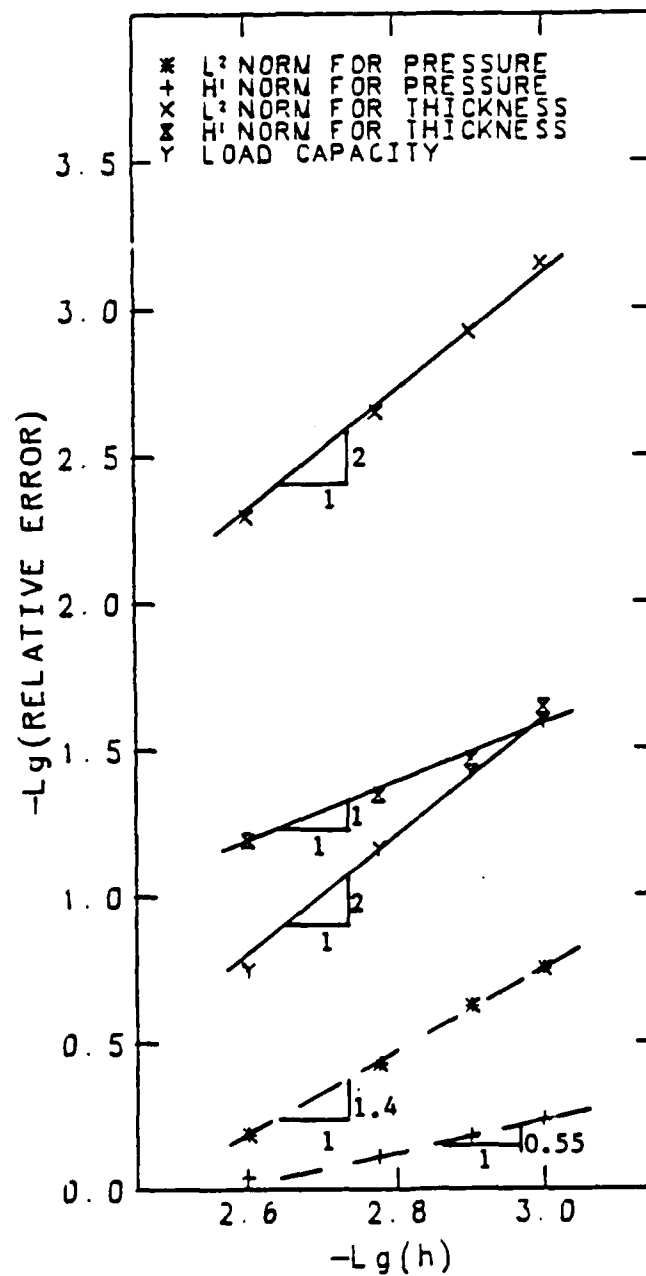


Fig. 6.12 Error Analysis (2-D Bilinear Elements, $H_0=0.5E-6$)
Point Contact ($W=0.142E-7$)

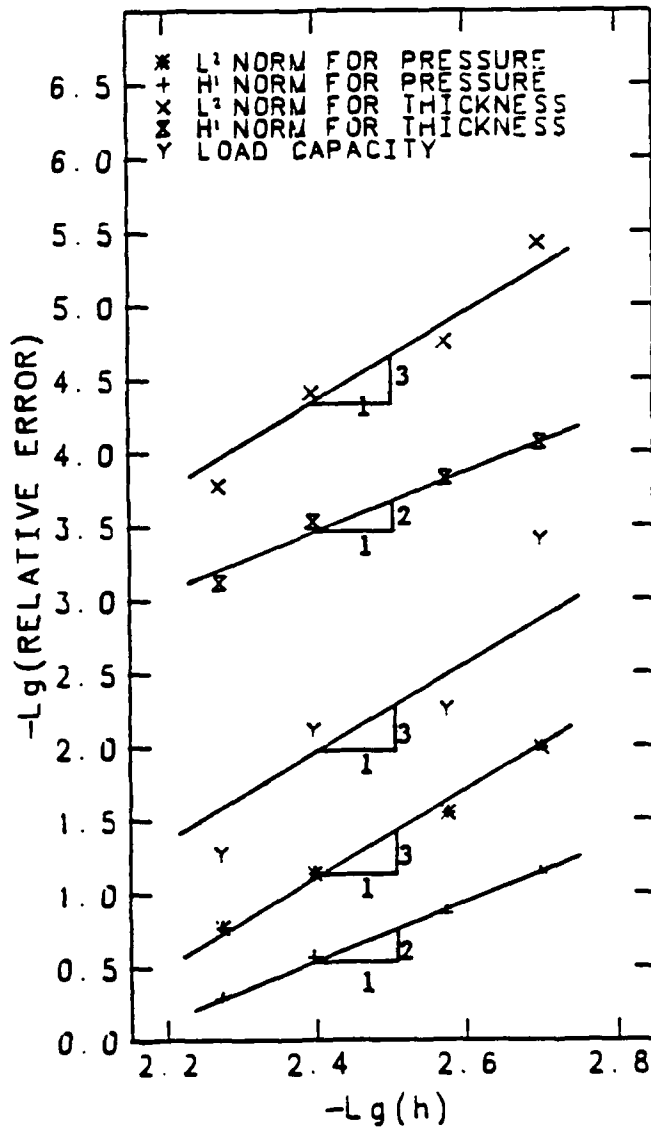


Fig. 6.13 Error Analysis (2-D Biquadratic Elements, $H_0=0.9E-5$)
Point Contact ($W=0.463E-8$)

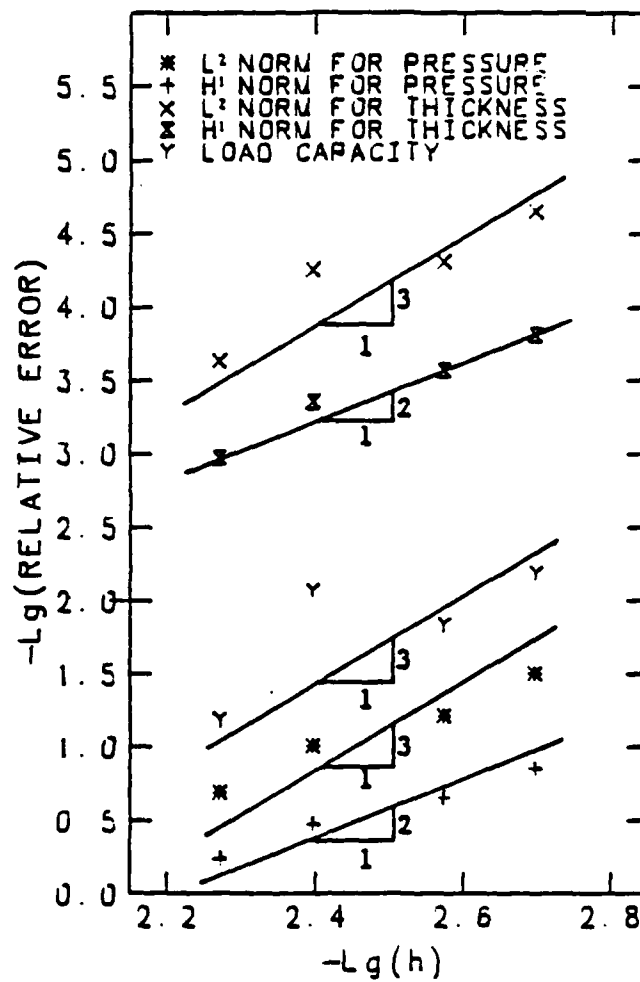


Fig. 6.14 Error Analysis (2-D Biquadratic Elements, $H_0=0.75E-5$)
Point Contact ($W=0.528E-8$)

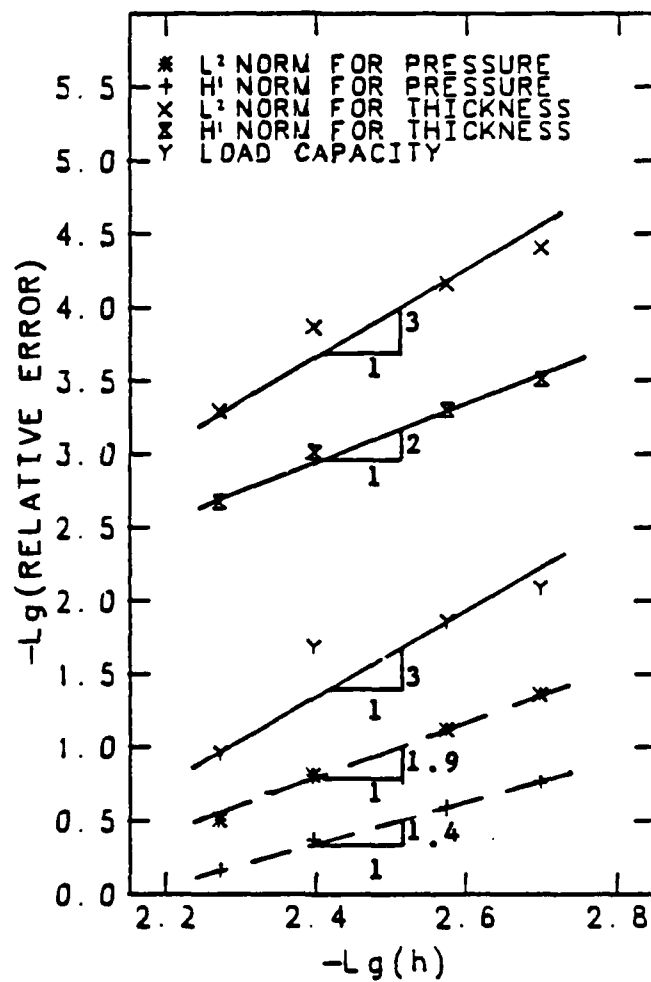


Fig. 6.15 Error Analysis (2-D Biquadratic Elements, $H_0=0.5E-5$)
Point Contact ($W=0.694E-8$)

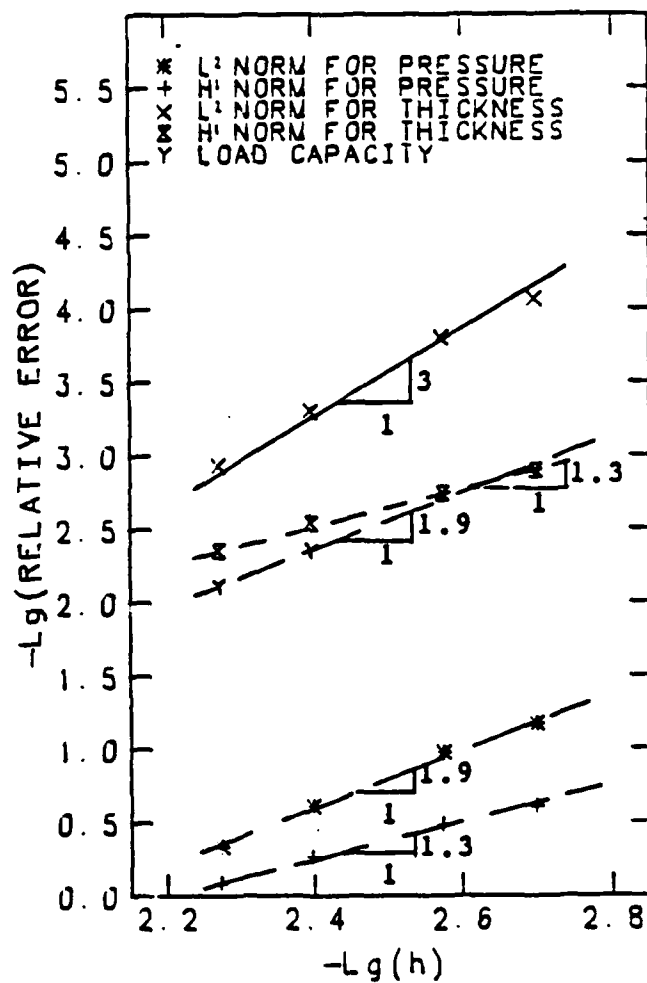


Fig. 6.16 Error Analysis (2-D Biquadratic Elements, $H_0=0.3E-5$)
Point Contact ($W=0.826E-8$)

To do so, we need an a-posteriori estimation to select the elements in which to perform the refinement. Meanwhile, some special techniques are implemented to match the refinement scheme.

6.5.1 A-POSTERIORI ERROR ESTIMATE.

In section 6.4 we derived the a-priori error estimate:

$$\|\text{error}\|_{H^1} \leq Ch^k |p_\epsilon|_{H^{k+1}}.$$

We shall use the number $h^k |p_\epsilon|_{H^{k+1}}$ as an error indicator for each element and pick the elements with larger indicators for refinement. The idea is to attempt to distribute the error evenly throughout the mesh.

For Lagrangian finite elements, used in this study, the functions interpolated with shape functions are continuous but may have jumps in the derivatives across the interelement boundaries. This fact must be considered in implementing the refinement procedure.

For one-dimensional linear elements (the $k=1$ case), the first derivative is constant in each element. In fact, it is a step function. When we calculate the second order seminorm, we obtain δ -functions. Thus, we can use the jumps of the first derivative on the interelement boundaries to estimate the local H^2 -seminorm of the solution:

$$\int (d^2p/dx^2)^2 dx = \sum_i ([dp/dx] |_{x_i})^2$$

For two-dimensional problems, the rectangular bilinear elements have the same feature, except that the jumps on the boundaries $x=x^i$ (or $y=y^j$) are still functions of y (or x respectively), so we may take the integrals of the jumps along the boundaries:

$$\begin{aligned} & \int [(\partial^2 p / \partial x^2)^2 + (\partial^2 p / \partial y^2)^2] d\Omega \\ &= \sum_i \int ([\partial p / \partial x] |_{x^i})^2 dy + \sum_j \int ([\partial p / \partial y] |_{y^j})^2 dx \end{aligned}$$

However, for the quadratic elements ($k=2$ case), the third order seminorm seems awkward to use in such computations. Hence we prefer to the advantage of the lower-order criterion employed for linear elements, but take into account the

contribution to the indicator of results defined interior to elements.

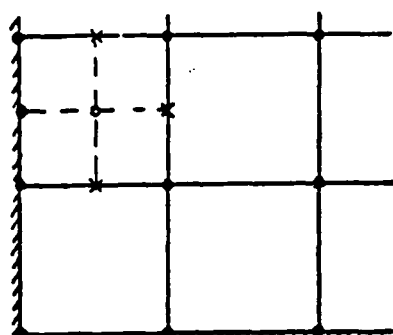
6.5.2 REFINEMENT SCHEME.

For one-dimensional problems, a simple refinement scheme just bisects some elements and increases the number of nodes. If the original problem formulated with a banded matrix, the refinement can still result in a banded matrix, after refinement and renumbering, with the same bandwidth.

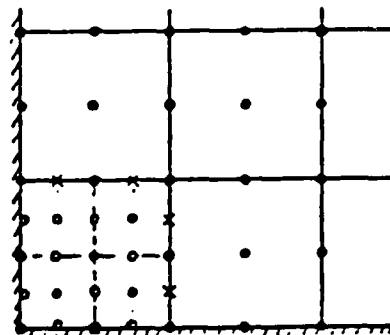
However, for two-dimensional problems, the development of an efficient renumbering strategy is not an easy task. Even if an optimal numbering process is implemented and a banded matrix results, the bandwidth may increase dramatically. Thus the band solvers are no longer suitable.

In this work, we use the iterative scheme based on the fundamental structure of finite element method that the matrix is assembled by the element matrix. So if we use any iterative solver, e.g. Lanczos (see [95]), which employs matrix-vector multiplications only, these multiplications being performed on the element level. Thus we only need store the information on the element level. This element-by-element scheme (see [47]) is very convenient for the refinement procedure. In particular, it is easy to coordinate with the H-split formulation in Section 5.3.

For simplicity, we use rectangular elements of the same order. When we refine an element into four elements, new nodes appear on the element boundaries. If the neighboring element is not to be refined, it will have extra nodes on its boundary. To maintain consistency, these new nodes should be constrained, as shown in Fig. 6.17(a), by an interpolation of the values on the neighboring element. Of course, if the new nodes are on the boundary of the domain of interest, they are not confined on the element, but are imposed by corresponding boundary conditions. If, in the next refinement, the neighboring element is refined, as shown in Fig. 6.17(b), the constraint on the node should be removed. Meanwhile, we employ the one-node rule, which means that on one side of a linear element at most one constrained node (a pair for quadratic elements) is allowed. Thus, when a second node on the side is to be constrained, the neighboring element should be refined instead, as shown in Fig. 6.17(c).



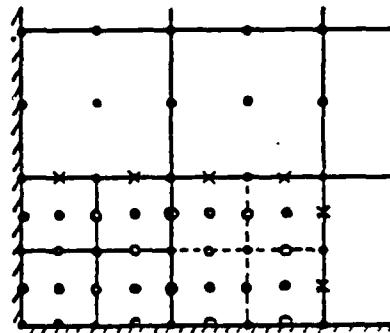
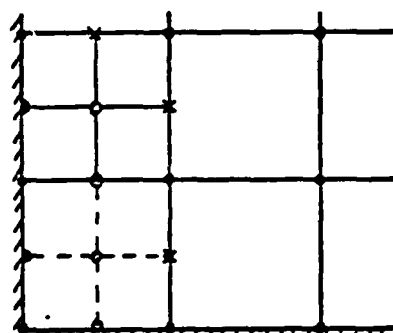
Bilinear Elements



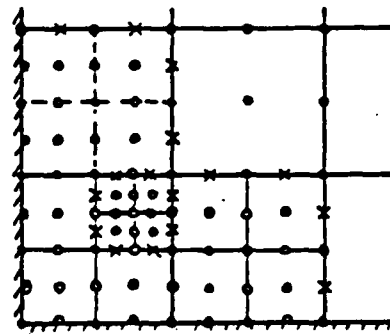
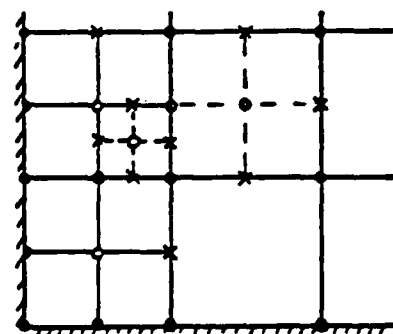
Biquadratic Elements

- Basic Nodes
- Added Active Nodes
- * Constrained Nodes

(a) Constrained Nodes in the First Refined Mesh



(b) Constrained Nodes in the Second Refined Mesh



(c) The Third and Subsequent Refinement

Fig. 6.17 Refinement of Rectangular Elements

6.5.3 APPLICATIONS TO LINE CONTACT PROBLEMS.

Figure 6.18 shows the pressure distributions solved on a series of adaptively refined linear-element meshes starting with a uniform 10-element mesh, For a light load case ($H_0=0.$, $W=0.304E-5$). It is shown that the solutions from the fourth (28 elements) and the fifth (33 elements) refined meshes are very close. In fact, the relative difference of pressure $\|p_4-p_5\|/\|p_5\|$ is 5.0 percent in the H^1 -norm and 0.75 percent in the L^2 -norm. The relative difference of load is 0.40×10^{-3} . For the film thickness, it is 0.96 percent in the H^1 -norm and 0.41×10^{-3} in the L^2 -norm. Figure 6.19 presents the refinement patterns.

Figure 6.20 shows the pressure distributions solved on a series of adaptively refined quadratic-element meshes starting with a uniform 5-element mesh. The relative difference of pressure between the solutions from the fourth (17 elements) and the fifth (20 elements) refined meshes is 2.0 percent in the H^1 -norm and 0.45 percent in the L^2 -norm. For film thickness, it is 0.37 percent in the H^1 -norm and 0.24×10^{-3} in the L^2 -norm. And the difference in load is 0.54×10^{-3} . Figure 6.21 presents the refinement patterns.

It is interesting to make a comparison with the solutions from uniform fine meshes. In Fig. 6.22, the pressure distributions solved on a set of uniform linear-element meshes are shown (referring the convergence shown in Fig. 6.2). Figure 6.23 presents the pressure distributions solved on uniform quadratic-element meshes (referring the convergence shown in Fig. 6.5). Figure 6.24(a) shows the comparison between the solutions from the fifth refined linear-element mesh (33 elements) and a uniform fine mesh (120 elements). The comparison for solutions with quadratic elements is shown in Fig. 6.24(b). Figure 6.24(c) compares the linear and quadratic adaptive solutions. It is observed that the adaptive solutions and those from fine uniform meshes are very close, and the adaptive solutions with linear elements and quadratic elements are also very close. These results thus show that the adaptive method is quite efficient and reliable in obtaining an accurate approximation.

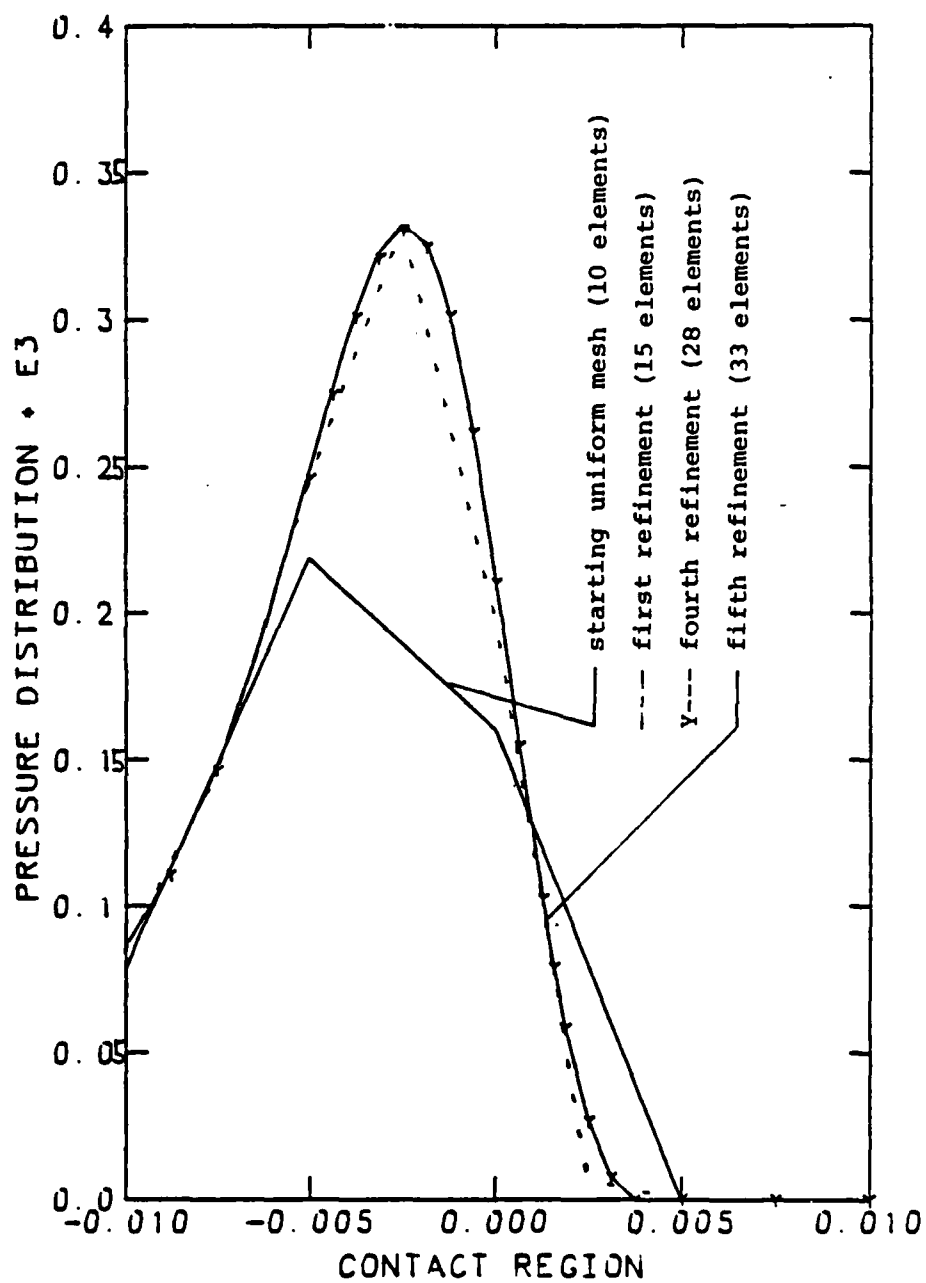


Fig. 6.18 Pressure Distributions
(1-D Adaptive Linear Elements, $H_0=0.0$)

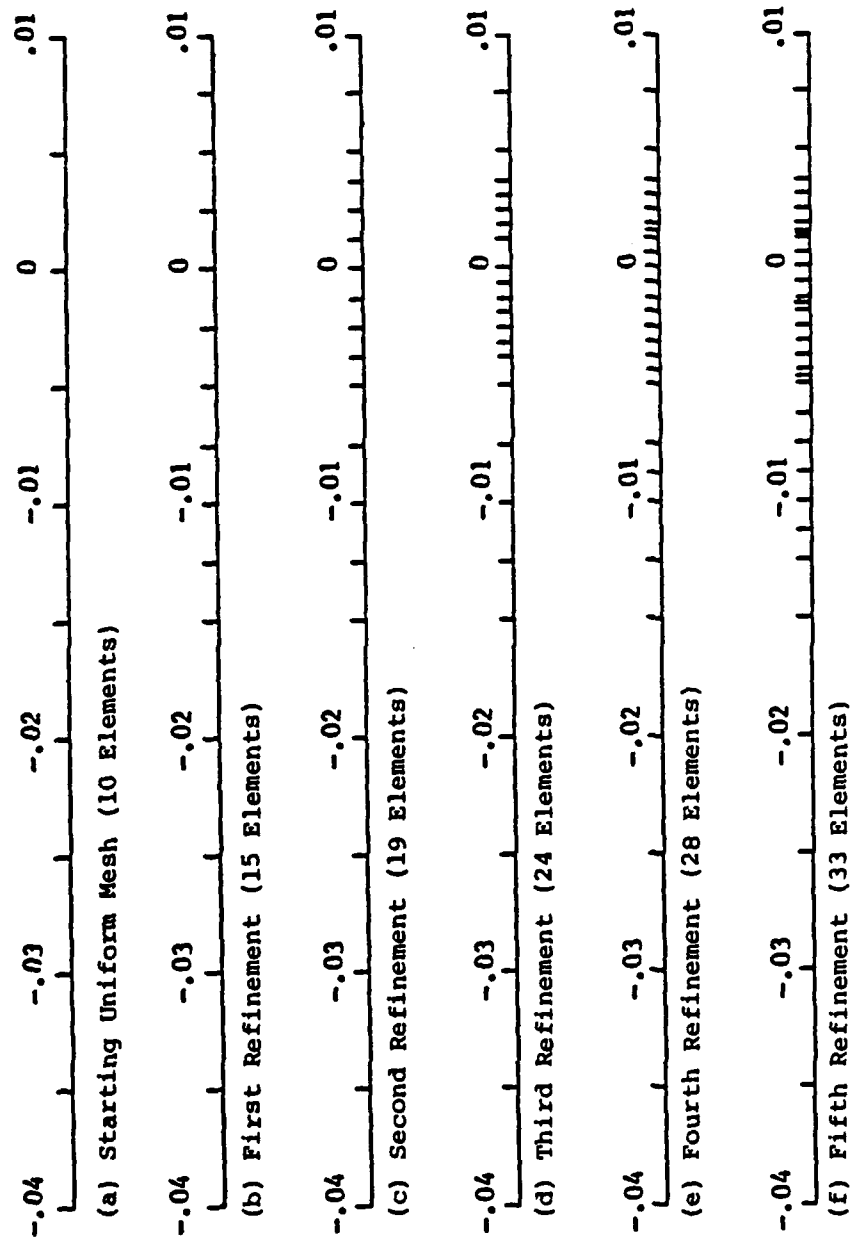


Fig. 6.19 Refinement Patterns for Fig. 6.18

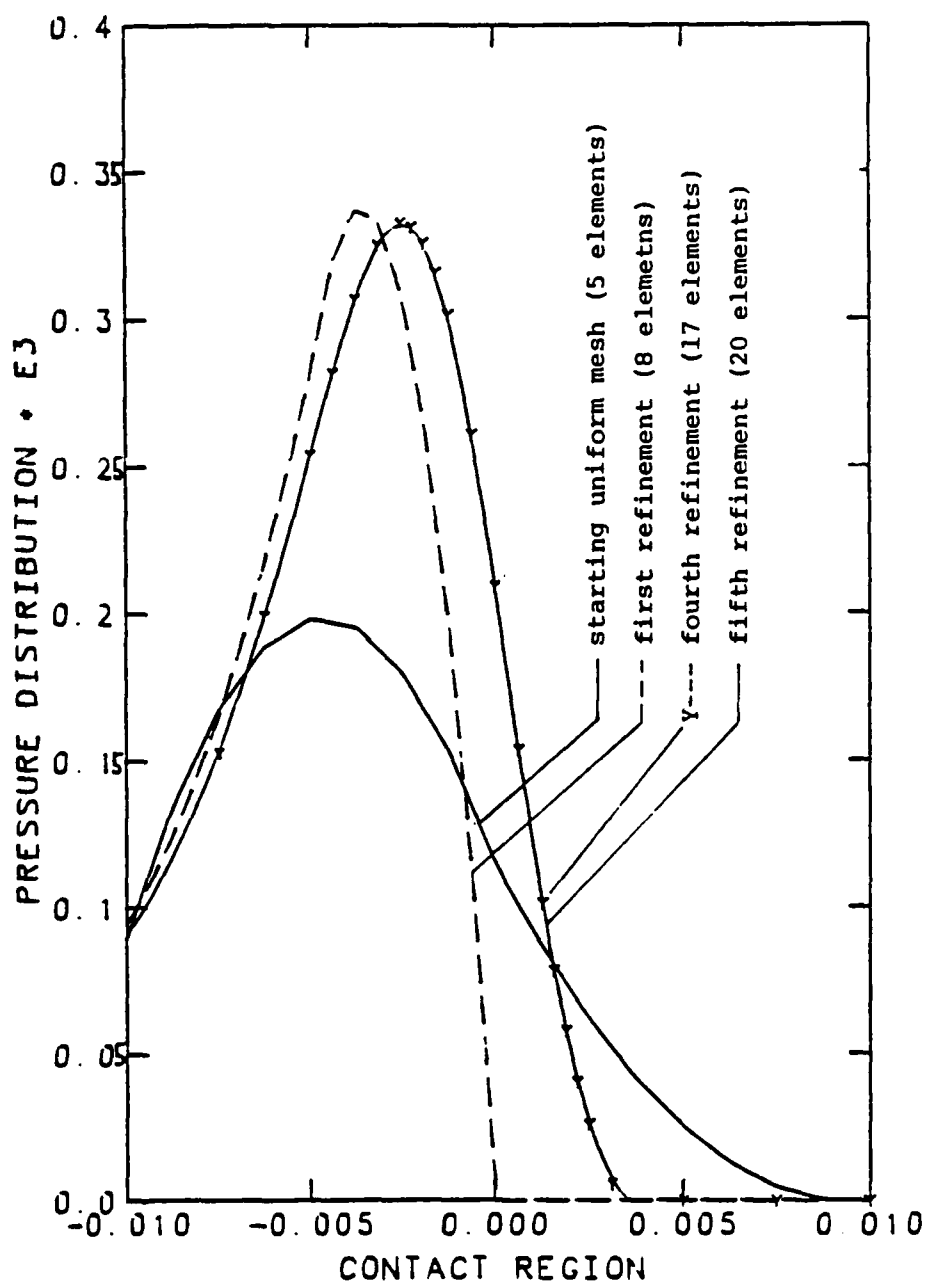


Fig. 6.20 Pressure Distributions
(1-D Adaptive Quadratic Elements, $H_0=0.0$)



(a) Starting Uniform Mesh (5 Elements)



(b) First Refinement (8 Elements)



(c) Second Refinement (11 Elements)



(d) Third Refinement (14 Elements)



(e) Fourth Refinement (17 Elements)



(f) Fifth Refinement (20 Elements)

Fig. 6.21 Refinement Patterns for Fig. 6.20

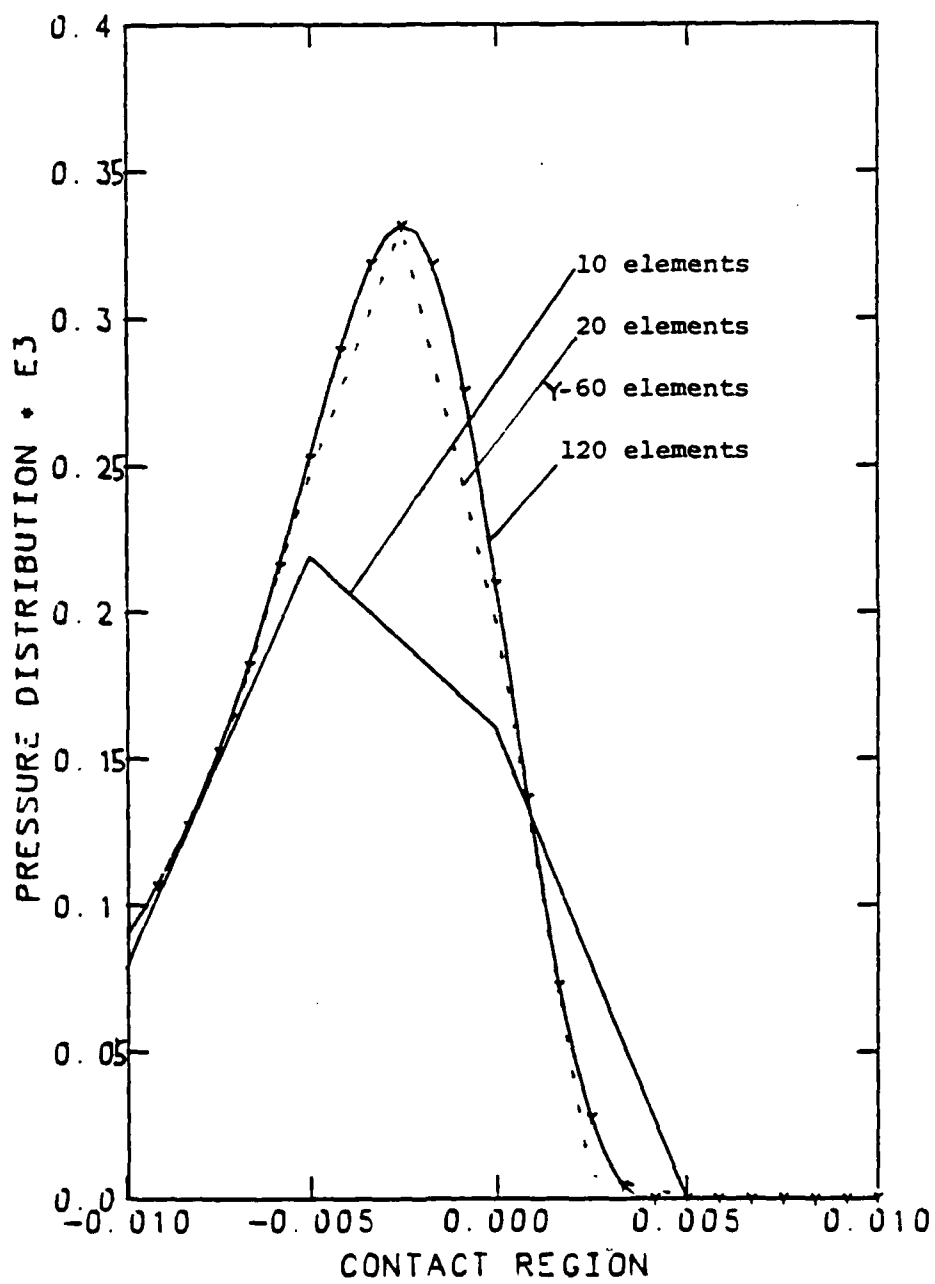


Fig. 6.22 Pressure Distributions
(1-D Uniform Linear Elements, $H_0=0.0$)

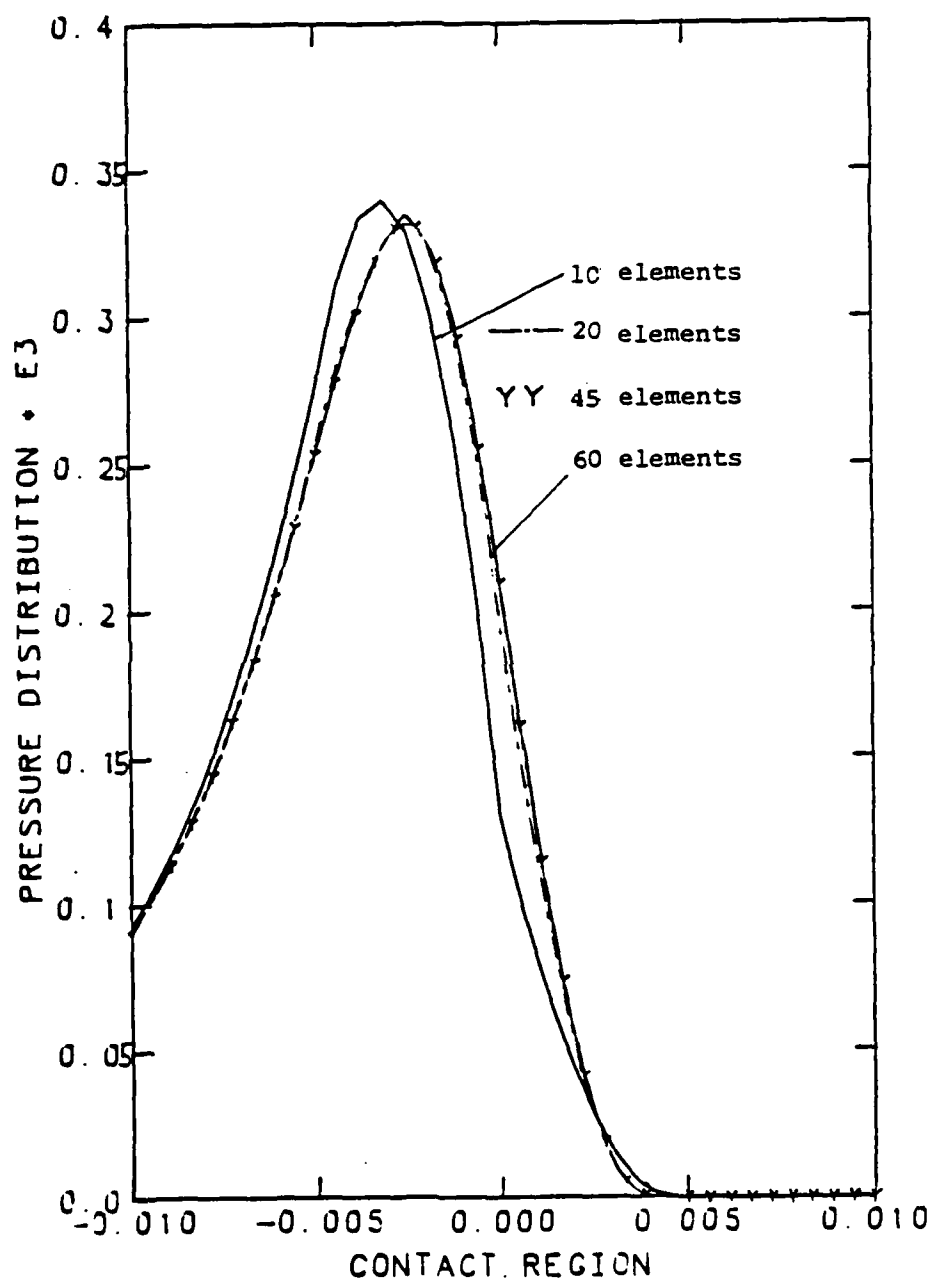
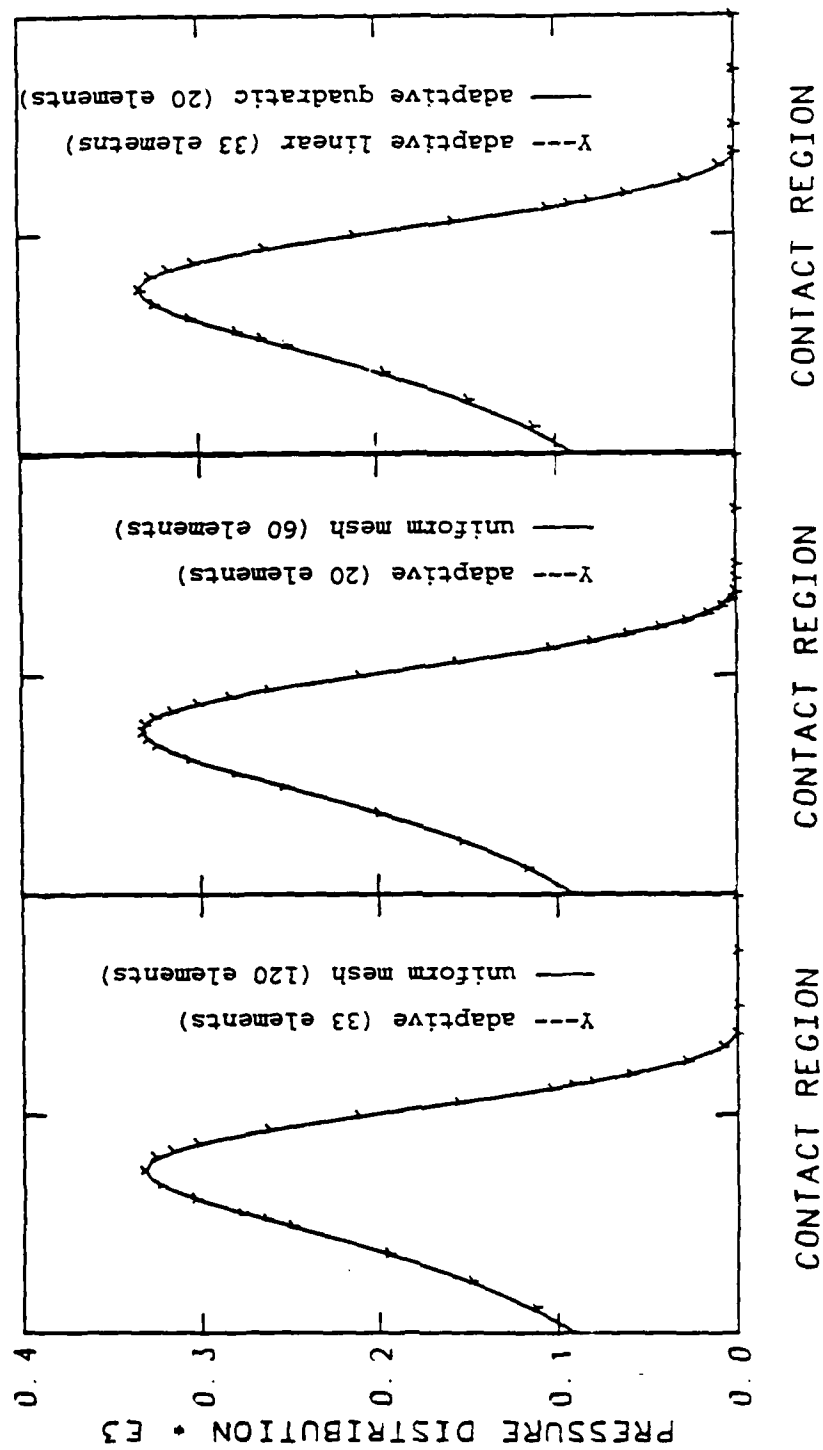


Fig. 6.23 Pressure Distributions
(1-D Uniform Quadratic Elements, $H_0=0.0$)



(a) Linear-Element Solutions (b) Quadratic-Element Solutions (c) Adaptive Solutions Linear/Quadratic

Fig. 6.24 Comparisons with Fine Uniform Meshes (1-D, $H_0=0.0$)

A second example is for a heavy load case ($H_0 = -0.16E-3$, $W = 0.233E-4$). Figure 6.25 shows the pressure distributions solved on a series of adaptive-refined linear-element meshes starting with a 10-element uniform mesh. The refinement patterns are shown in Fig. 6.26. The relative difference between the pressure solved on the fifth (35 elements) and the sixth (40 elements) refined meshes is 8.9 percent in the H^1 -norm and 0.49 percent in the L^2 -norm. The difference in load is 0.52×10^{-5} . And the difference in the film thickness is 0.35 percent in the H^1 -norm and 0.48×10^{-4} in the L^2 -norm.

The computation for quadratic elements is shown in Fig. 6.27. Figure 6.28 shows the refinement patterns. In the process, the scheme at first appeared to be divergent at the load level, the first two refinements are performed at a lower loading level. Then we increased the load again and continued the refinement procedures. The relative difference of pressure between the fourth (44 elements) and fifth (49 elements) refinement solutions is 3.1 percent in the H^1 -norm and 0.17 percent in the L^2 -norm. The difference in load is 0.28×10^{-5} . And for the film thickness, 0.25 percent in the H^1 -norm, and 0.22×10^{-4} in the L^2 -norm.

Figure 6.29 exhibits the comparisons of adaptive solutions with fine uniform mesh solutions. Figure 6.30 shows the film shape near contact center.

These experiments and the comparisons show that adaptive methods can be very effective for these types of nonlinear problems.

6.5.4 APPLICATION TO POINT CONTACT PROBLEM.

For a light load case ($H_0 = 0.5E-5$), computed with bilinear elements, after three refinements the adaptive solutions become quite close to each other. The pressure profiles at $y=0$, of the adaptive solutions are shown in Fig. 6.31. Figure 6.32 exhibits the refined meshes generated by the adaptive method.

The fifth (369 elements) and the sixth (456 elements) adaptive solutions have a difference in pressure about 19 percent in the H^1 -norm and 3.2 percent in the L^2 -norm. The difference in the film thickness is 1.7 percent in the H^1 -norm and 0.11 percent in the L^2 -norm. The difference in load is 1.5 percent.

Figure 6.33 shows the result for biquadratic-elements. Finally, the fifth

(237 elements) and sixth (285 elements) refinement solutions produces a difference in pressure about 29 percent in the H^1 -norm and 3.0 percent in the L^2 -norm. The difference in the film thickness is 0.16×10^{-3} in the H^1 -norm and 0.17×10^{-4} in the L^2 -norm. The difference in load is 0.18 percent. Figure 6.34 shows the refinement mesh patterns. Comparisons of adaptive solutions with the solutions from fine uniform meshes are shown in Fig. 6.35.

These experiments on the light load cases encourage and suggest that further studies of adaptive techniques for these types of problems may be warranted.

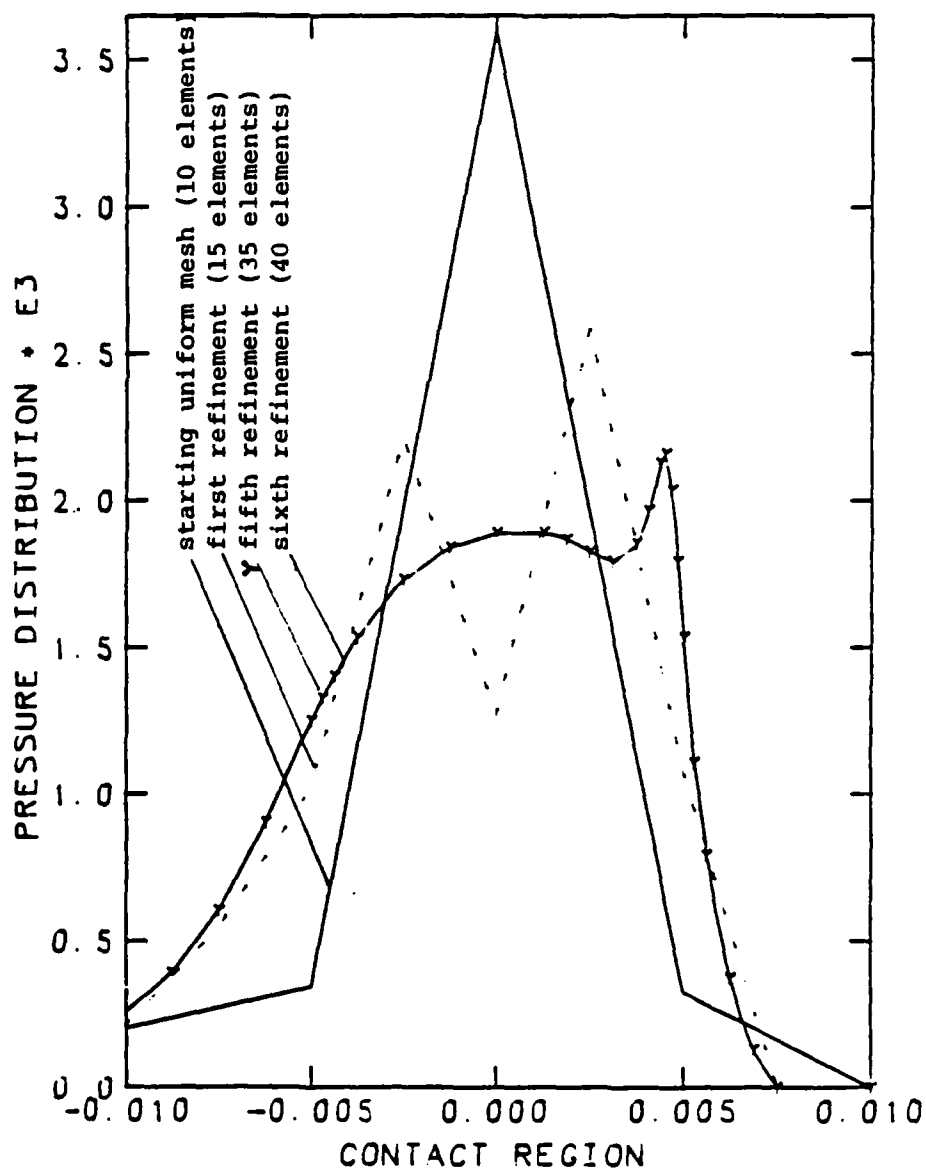


Fig. 6.25 Pressure Distributions
(1-D Adaptive Linear Elements, $H_0 = -0.16E-3$)

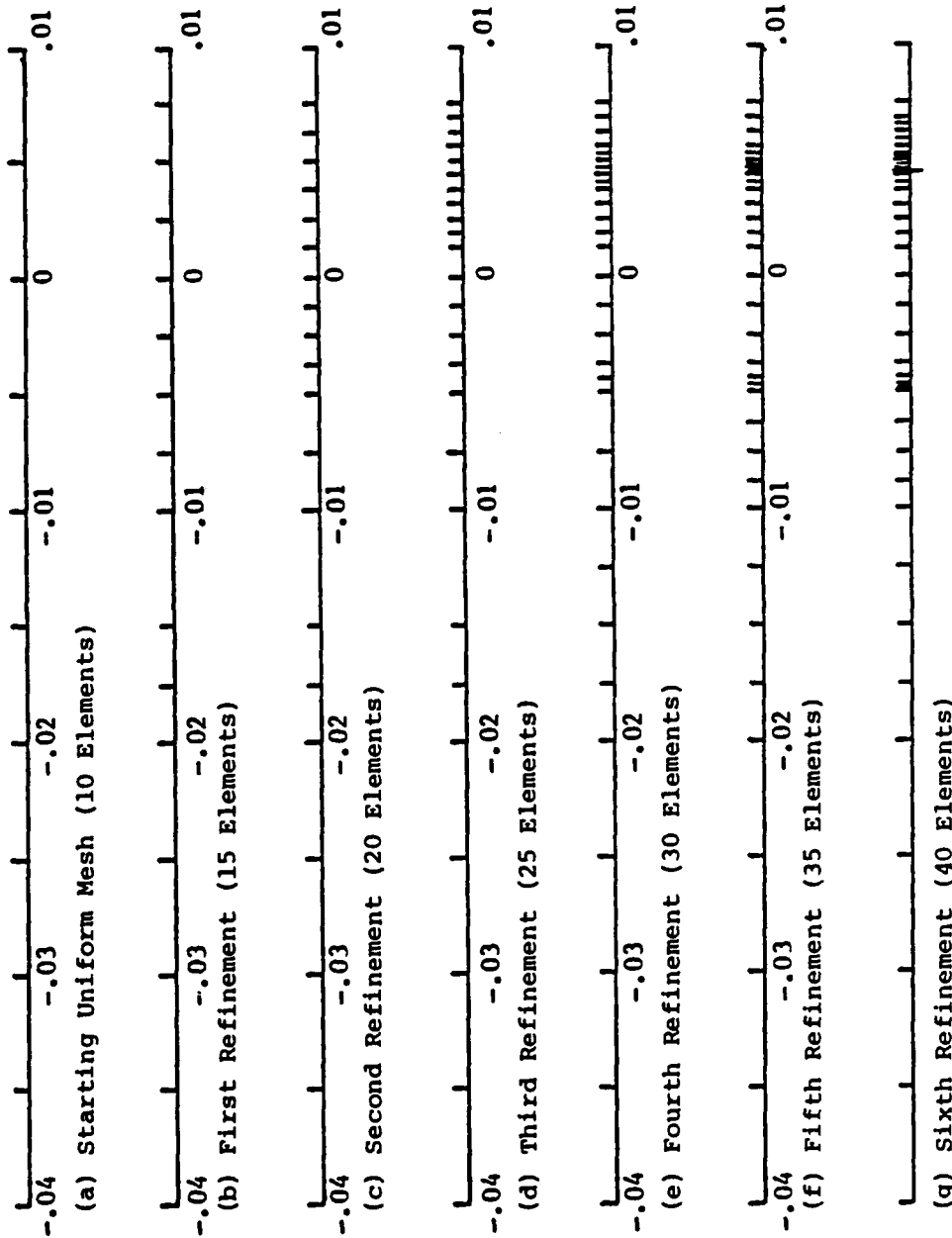


Fig. 6.26 Refinement Patterns for Fig. 6.25

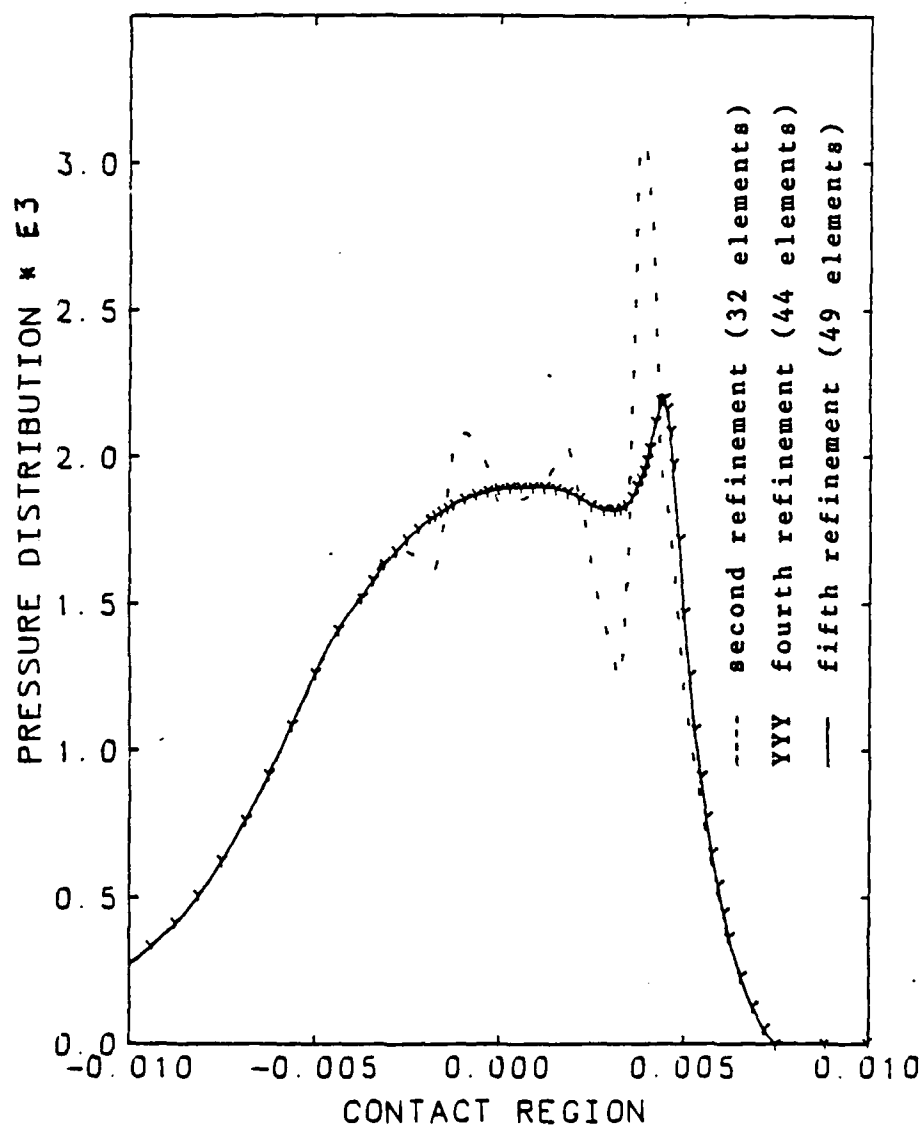


Fig. 6.27 Pressure Distributions
(2-D Adaptive Quadratic Elements,
 $H_0 = -0.16E-3$)

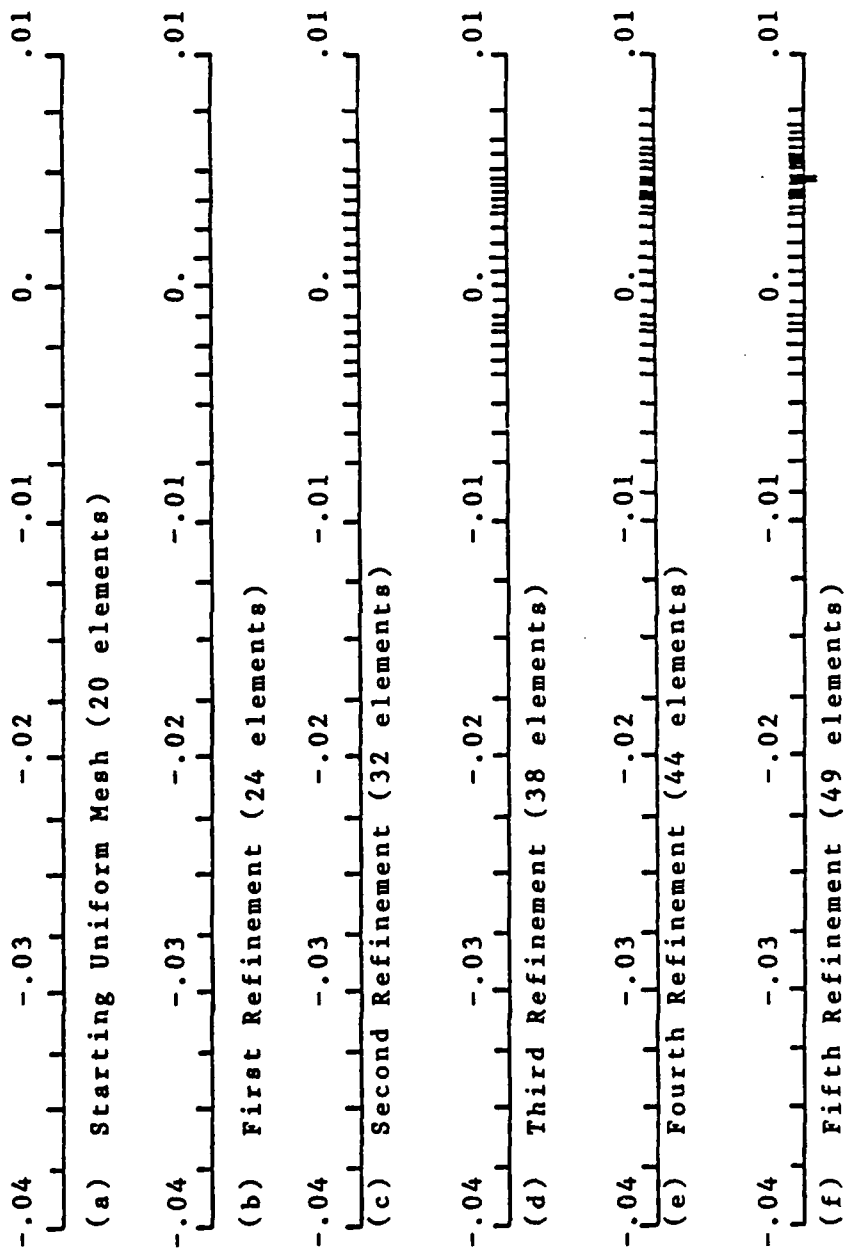
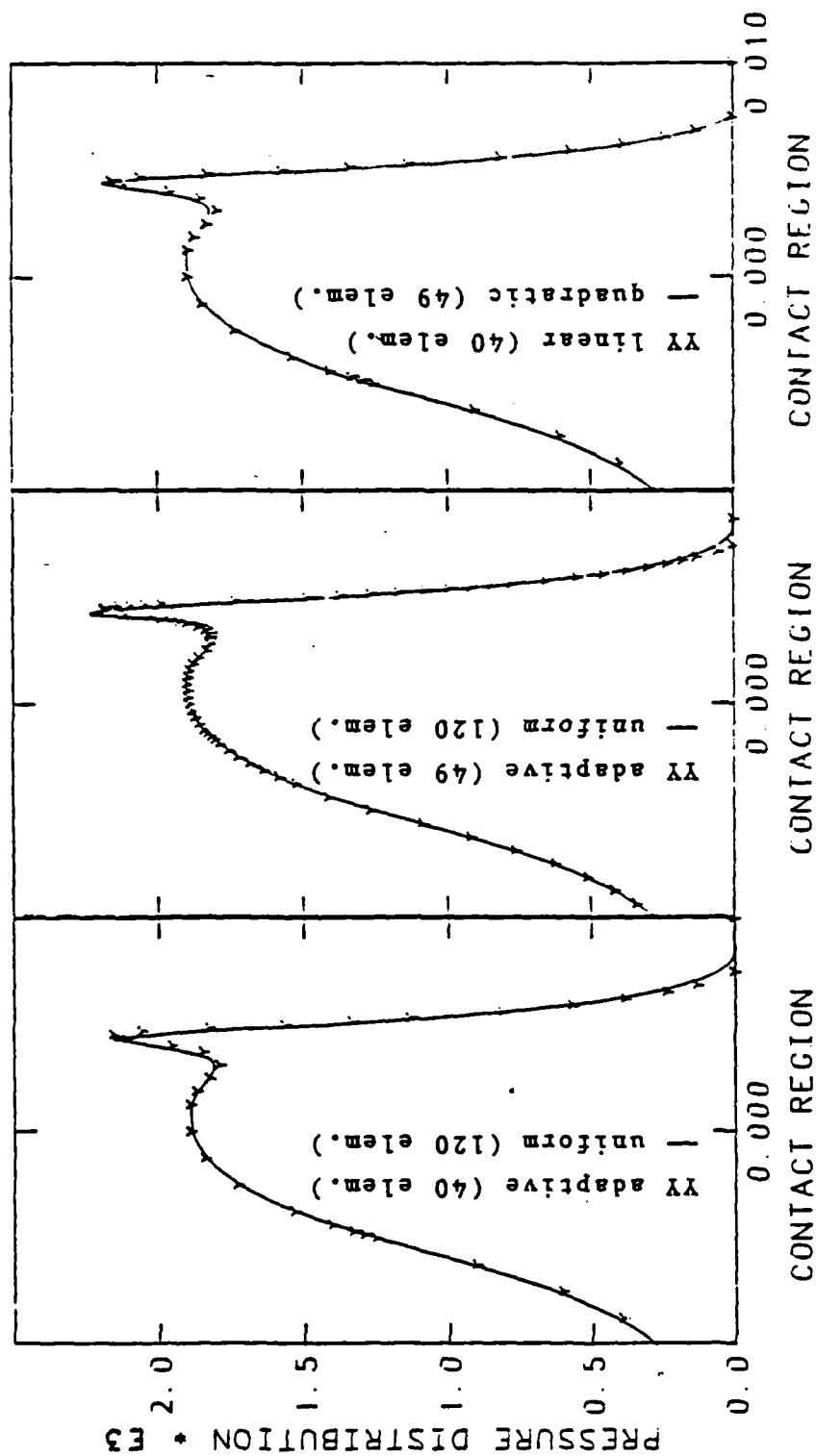


Fig. 6.28 Refinement Patterns for Fig. 6.27



(a) Linear Elements (b) Quadratic Elements (c) Adaptive Solutions

Fig. 6.29 Comparison with Fine Uniform Mesh (1-D, $H_0 = -0.16E-3$)

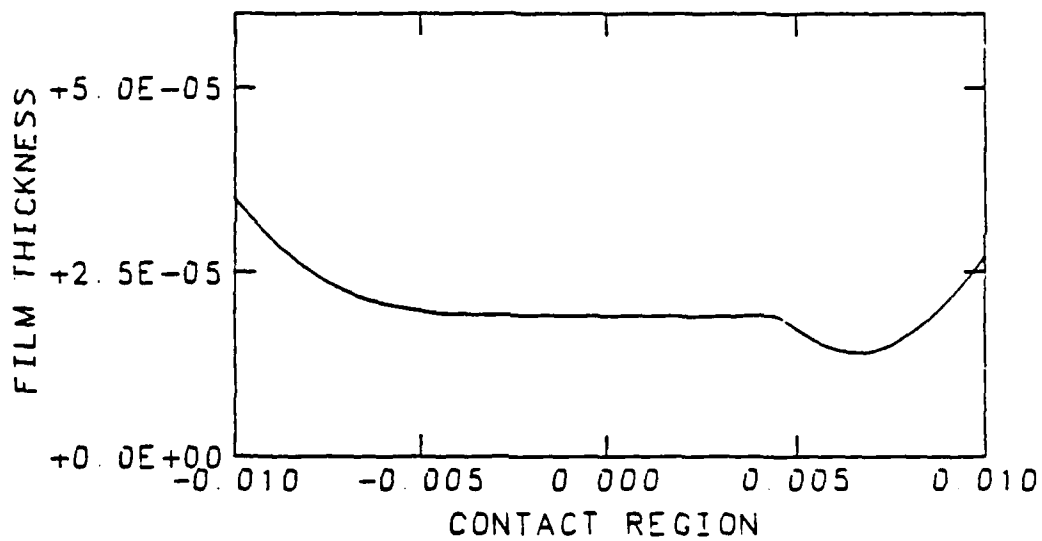


Fig. 6.30 Film Shape near Contact Center
 $(1-D, H_0 = -0.16E-3)$

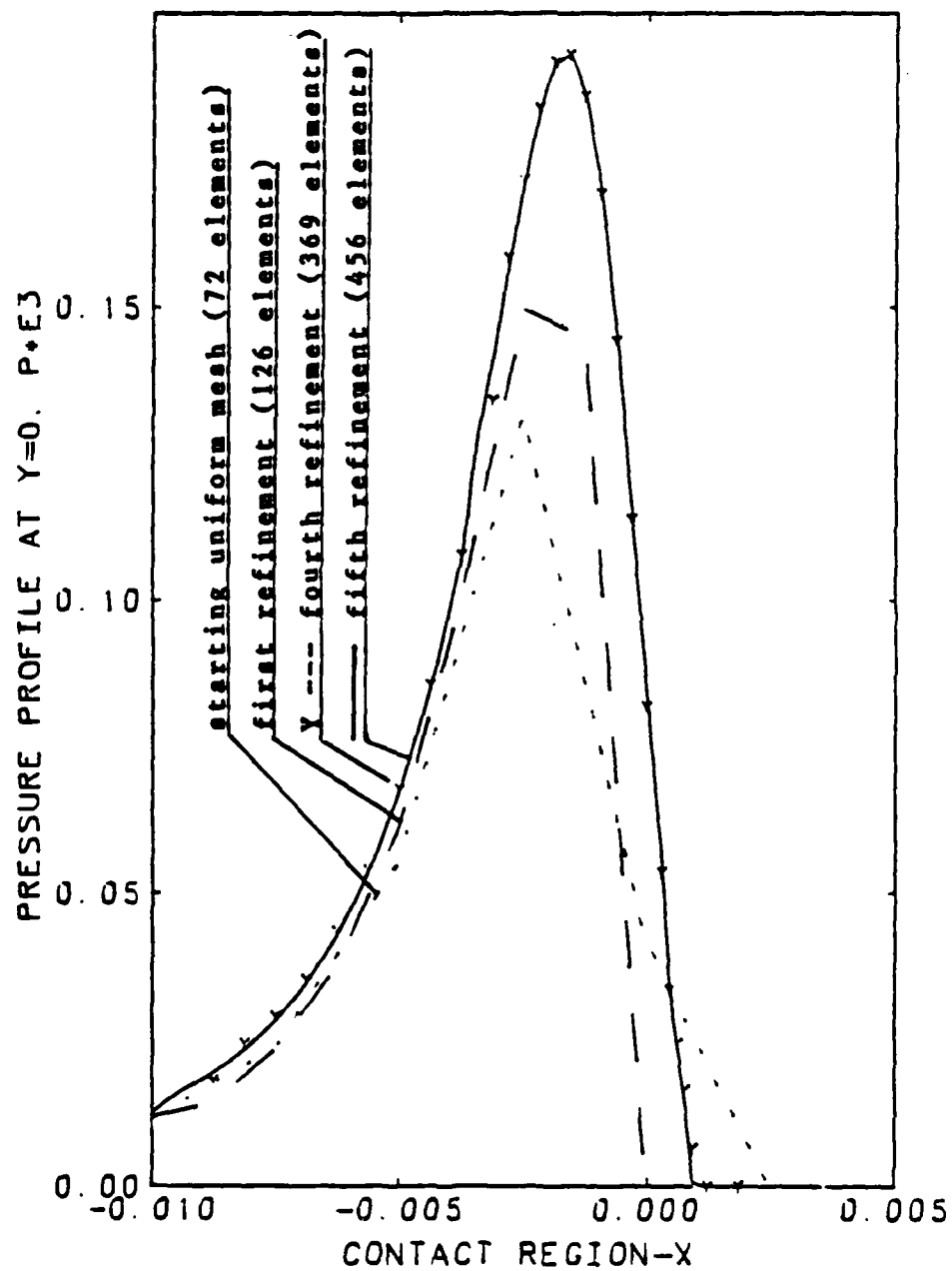
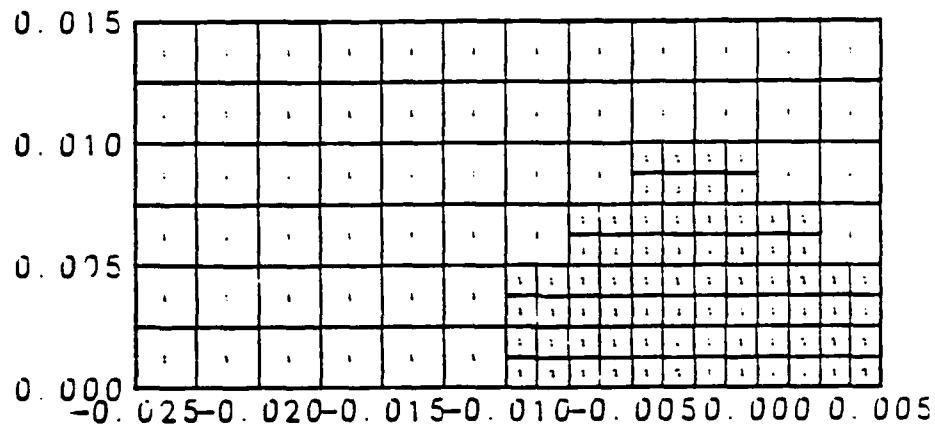
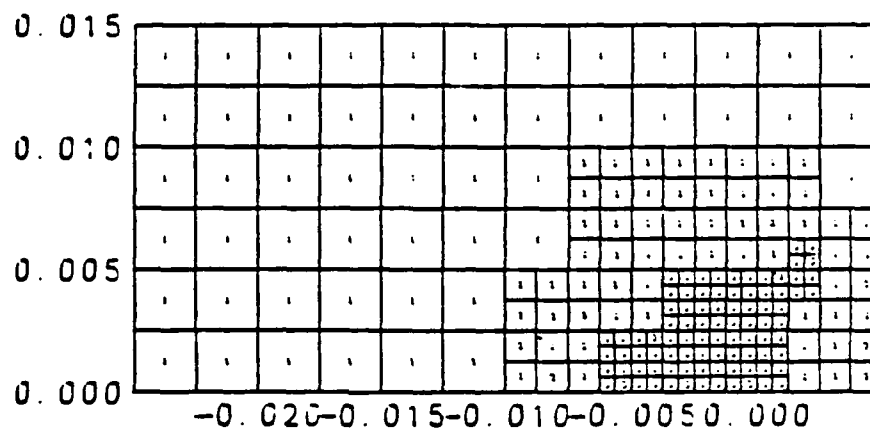


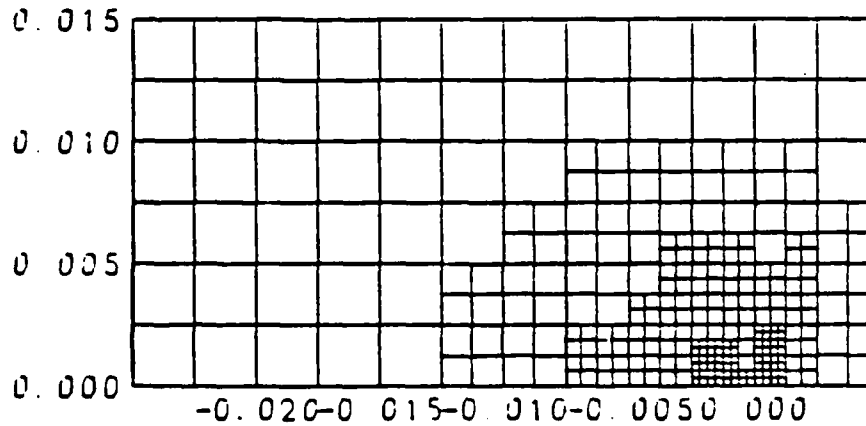
Fig. 6.31 Pressure Profiles of Adaptive Solutions at $y=0$
(2-D Bilinear Elements, $H_0=0.5E-5$)



(a) First Refinement (126 elements)



(b) Second Refinement (201 elements)



(c) Third Refinement (291 elements)

Fig. 6.32 Refinement Patterns for Fig. 6.31

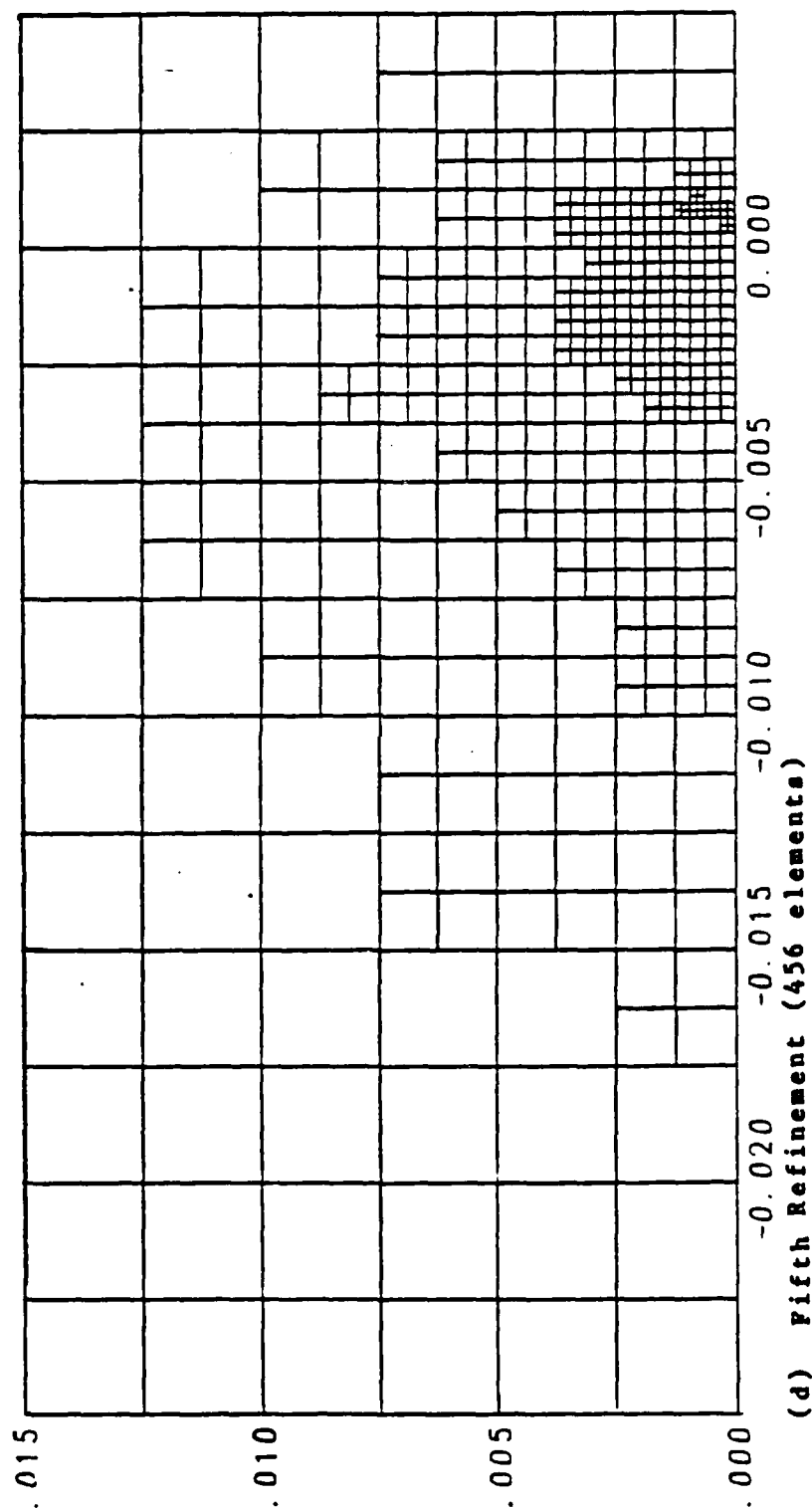


Fig. 6.32 Refinement Patterns for Fig. 6.31

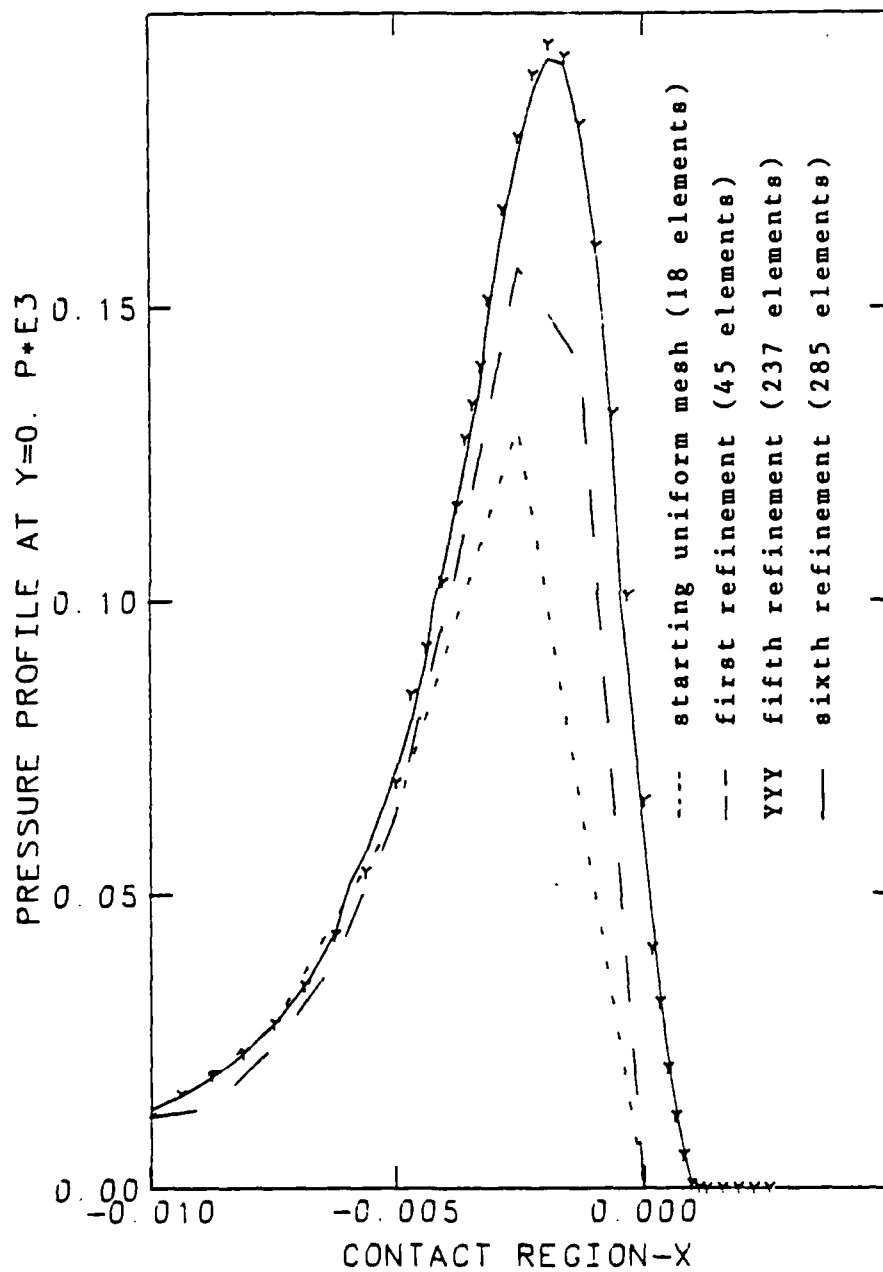
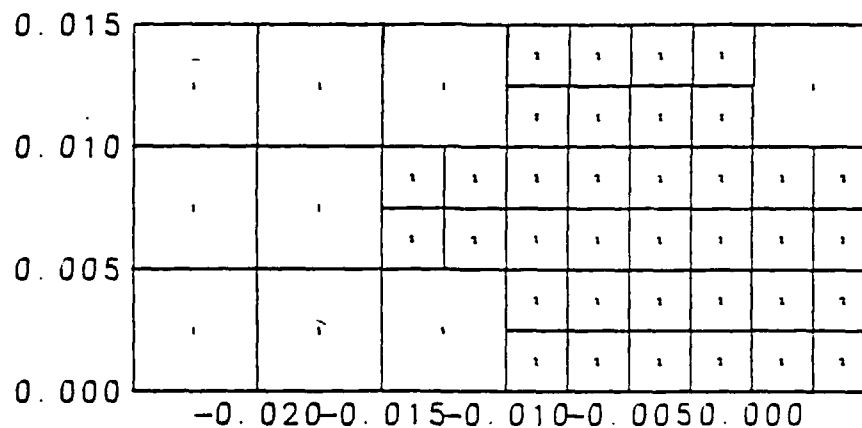
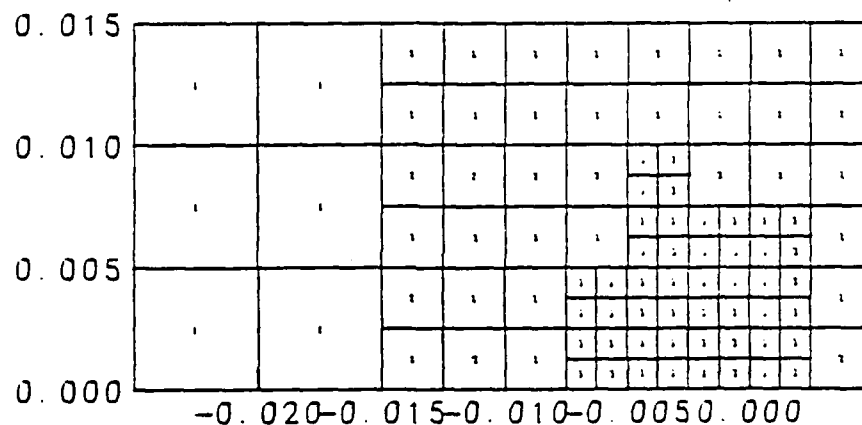


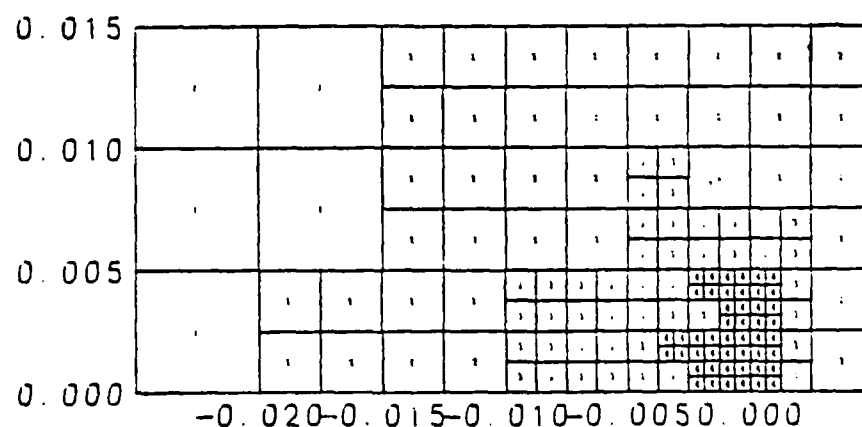
Fig. 6.33 Pressure Profiles of Adaptive Solutions at $y=0$ (2-D Biquadratic Elements, $H_0=0.5E-5$)



(a) First refinement (45 elements)

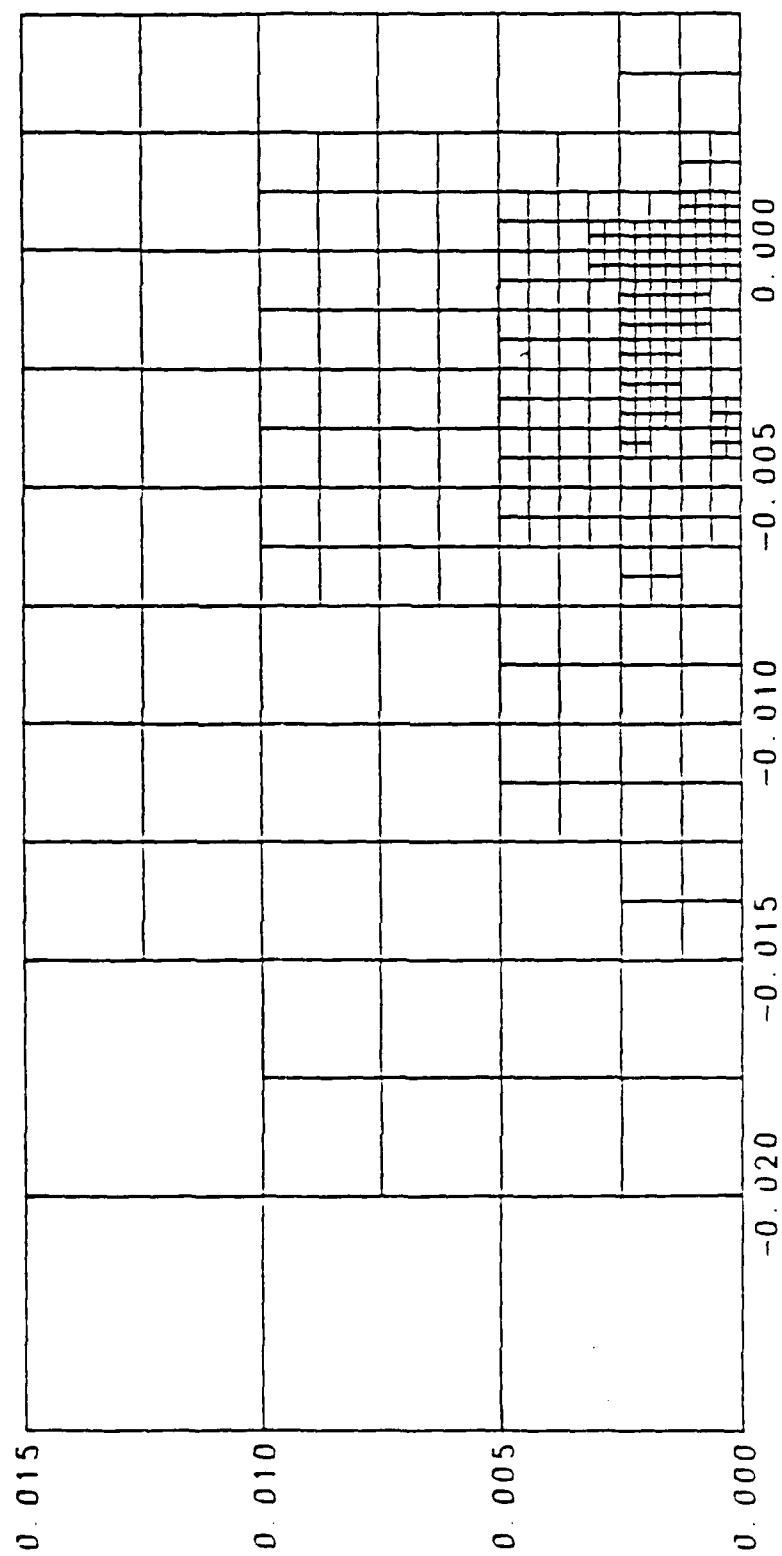


(b) Second refinement (90 elements)



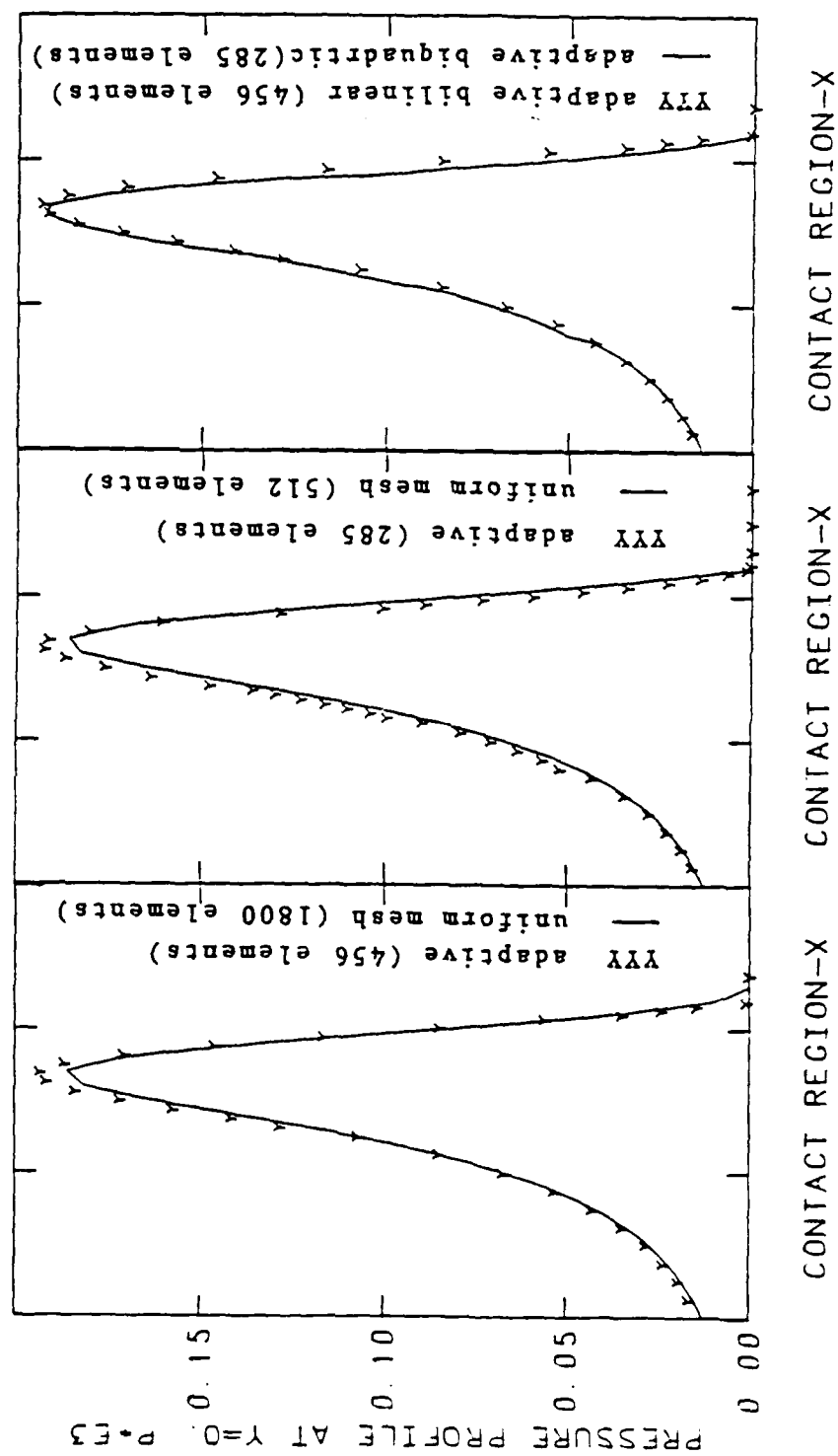
(c) Third refinement (144 elements)

Fig. 6.34 Refinement Patterns for Fig. 6.33



(d) Sixth refinement (285 elements)

Fig. 6.34 Refinement Patterns for Fig. 6.33



(a) Bilinear elements (b) Biquadratic elements (c) Adaptive solutions

Fig. 6.35 Comparison with Fine Uniform Mesh (2-D, $H_0=0.5E-5$)

7. CONCLUDING REMARKS

7.1 SUMMARY.

The elastohydrodynamic lubrication characterized by the highly nonlinear Reynolds–Hertz equations belongs to a class of free boundary problems. Formally, it may be classified as a problem governed by second order differential–integral equations. In this work, mathematical and numerical analyses have been performed in order to study the fundamental behavior of the equations and to demonstrate numerical schemes for solving practical problems in lubrication. The following steps were taken to accomplish these objectives:

(1) First, a variational inequality was formulated to provide a correct framework for mathematical analysis. The formulation led to a highly nonlinear variational inequality.

(2) The operator in the governing variational formulation was studied in some detail. It was proved that the operator is bounded, coercive, pseudomonotone, and continuous. These properties guarantee the existence of solutions to the variational inequality, and that the solutions are weak solutions of the Reynolds–Hertz equations. Under some circumstances when the bearing loads are light, the solution is unique.

(3) A penalty method was introduced to regularize the variational inequality with the release of the constraint on the free boundary. The penalized formulation of the problem provides a foundation for both mathematical analyses and numerical schemes.

The existence of solutions to the penalty problem follows as a direct consequence of the properties of the governing pseudomonotone operator. The convergence of penalty method was proved so that the penalty formulation was shown to be a valid basis for studying these classes of problems.

(4) Detailed studies of the one-dimensional line contact problem led to the

regularity of solutions through the analysis of the penalized formulation. As a consequence, the so-called Reynolds boundary flux condition at the free boundary was shown to be satisfied by the solutions. It was shown that there exist classical solutions to the line contact problem, which have smooth pressure distributions.

(5) Analyses based on the theory of Galerkin approximations in finite-dimensional spaces established a framework of the convergence relations among the penalized problem, the variational inequality and their approximations. A-priori error estimates were derived for finite element approximations, which predict that in certain cases of light loads, the pressure approximation exhibits first-order convergence with the linear elements and second-order convergence with quadratic elements, when measured in the H^1 -norm. For line contact problems, the same convergence rate for the film thickness h as for the pressure was also predicted.

(6) The finite element method was used for solving the penalized problems. Numerical results showed that the penalty method is excellent in locating the free boundary. The negative part of the pressure is always very small when a suitable penalty parameter ϵ is set, and reduces stably as the penalty number $1/\epsilon$ is increased. It was thus shown that the penalty method is applicable to obtain the approximate solutions for both the line contact and point contact problems.

Solutions of line contact problems with quadratic elements exhibit smooth pressure distributions. This fact is consistent with the regularity theorem for one-dimensional solutions.

(7) Error analysis of finite element solutions was carried out. Numerical experiments confirm the theoretical predictions on the rates of convergence. For line contact problems, the pressure and the film thickness exhibit first-order convergence in the H^1 -norm when computed with linear elements, and second-order convergence with quadratic elements. For point contact problems, the same behavior in convergence rates are also observed, but the valid range of loads is not as wide as in the case of line contact. A decrease in the convergence rate is observed in cases with heavier loads. One order higher convergence for the pressure and the film thickness measured in the L^2 -norm is also observed.

(8) Adaptive method was applied to the finite element solutions. For a light load case of line contact, adaptive solutions are in very good agreement with the fine uniform mesh solutions both with the linear elements and the quadratic elements. Also the adaptive solutions with linear elements and quadratic elements are very close. Adaptive solution for a heavy load case of line contact was also shown to converge and to be in good agreement with the fine uniform mesh solution, with linear elements. Application to the point contact problem with light load was performed with bilinear elements and the adaptive solutions also converged. Adaptive method was thus shown to be efficient and applicable to the highly nonlinear elastohydrodynamic lubrication problems.

7.2 RECOMMENDATIONS FOR FUTURE WORK.

(1) Based on this work, we can apply the penalty formulation to many other special cases in lubrication engineering, e.g., dynamically loaded elastohydrodynamic lubrication, thermal problems, gas bearings, and problems with more general non-Newtonian lubricants. The application to any kind of lubrication problem with an undetermined cavitation boundary might be direct and easy to implement. However, the theoretical analysis for time-dependent, temperature-dependent problems, and for non-Newtonian lubricants with highly nonlinear pressure-dependent viscosity should be significantly more difficult.

(2) Study is needed on the effects of numerical approximate quadratures used in the finite element formulations on numerical accuracy and convergence.

(3) Much additional work is needed on algorithm development, particularly on the development of schemes which use less storage and computer time but maintain the necessary accuracy. For example, domain-decomposition or block-relaxation techniques and adaptive refinement schemes deserve study.

(4) The following is a list of some difficult but significant open problems, encountered during this work, for which additional study is warranted.

(a) To prove or disprove the uniqueness of solutions of the variational inequality (3.3) and the penalty problem (3.27).

Physically, under certain conditions, the solution must be unique. In this

study we are only able to guarantee uniqueness under restriction to light loads via condition (4.33). Uniqueness, of course, is very important to the analysis, and a more delicate study is needed if any counterexample can be found in which bifurcation phenomenon occurs, then the study of the criterion, classification and behavior of bifurcation will form a new and difficult area of research which may be very significant to lubrication engineering.

(b) The regularity of solutions to point contact problems.

The establishment of regularity results for two-dimensional solutions is very difficult. It will involve the study of the integral $\int f(\xi)/r(x,\xi)d\Omega$ on a two-dimensional domain, with the singular kernel $r^{-1}(x,\xi)$. In the present study, the function f , the pressure, is known only as a H^1 -function.

(c) For two-dimensional, point contact problems under heavy loads, Fig. 7.1 shows the profiles of pressure distributions solved on a mesh with 560 bilinear elements. It is shown that when we increase load (reduce the parameter H_0), at first the change of pressure distribution exhibits some behavior similar to line contact problems (refer to Figs. 5.2 and 5.5). But when the load reaches a certain critical level, the pressure profile takes unexpected wiggles (such as the case of the first refinement in Fig. 6.25 or the second refinement in Fig. 6.27, for line contact) and this behavior may not resemble at all the true solution. A similar behavior is observed for quadratic elements and for some other meshes with more elements. However, the computer storage requirements preclude the use of finer meshes. Hence, the problem of systematically producing stable and accurate approximate solutions for heavy load cases of point contact remains open.

(d) In Reynolds' equation, physically the coefficient of the first derivative term is much larger than that of the second order terms. Particularly, under heavy loads, this situation may lead to numerical difficulties and numerical schemes may be unstable on all but very fine meshes.

For problems of this type, Petrov-Galerkin type variational formulations, upwind schemes, artificial viscosity, etc. have been used to improve the solutions for certain classes of convection-dominated problems, which have similar dominant

first-order term.

(e) When an adaptive method was applied to the point contact problems, the Galerkin formulation gave some unexpected wiggles in the pressure distribution. Figure 7.2 shows an example of such oscillation encountered in the third refinement. However, the Petrov-Galerkin formulation leads to a smooth solution, as shown in Figure 6.28. On the other hand, our adaptive refinement method appears to be unstable for heavy loads and additional study of this case is needed.

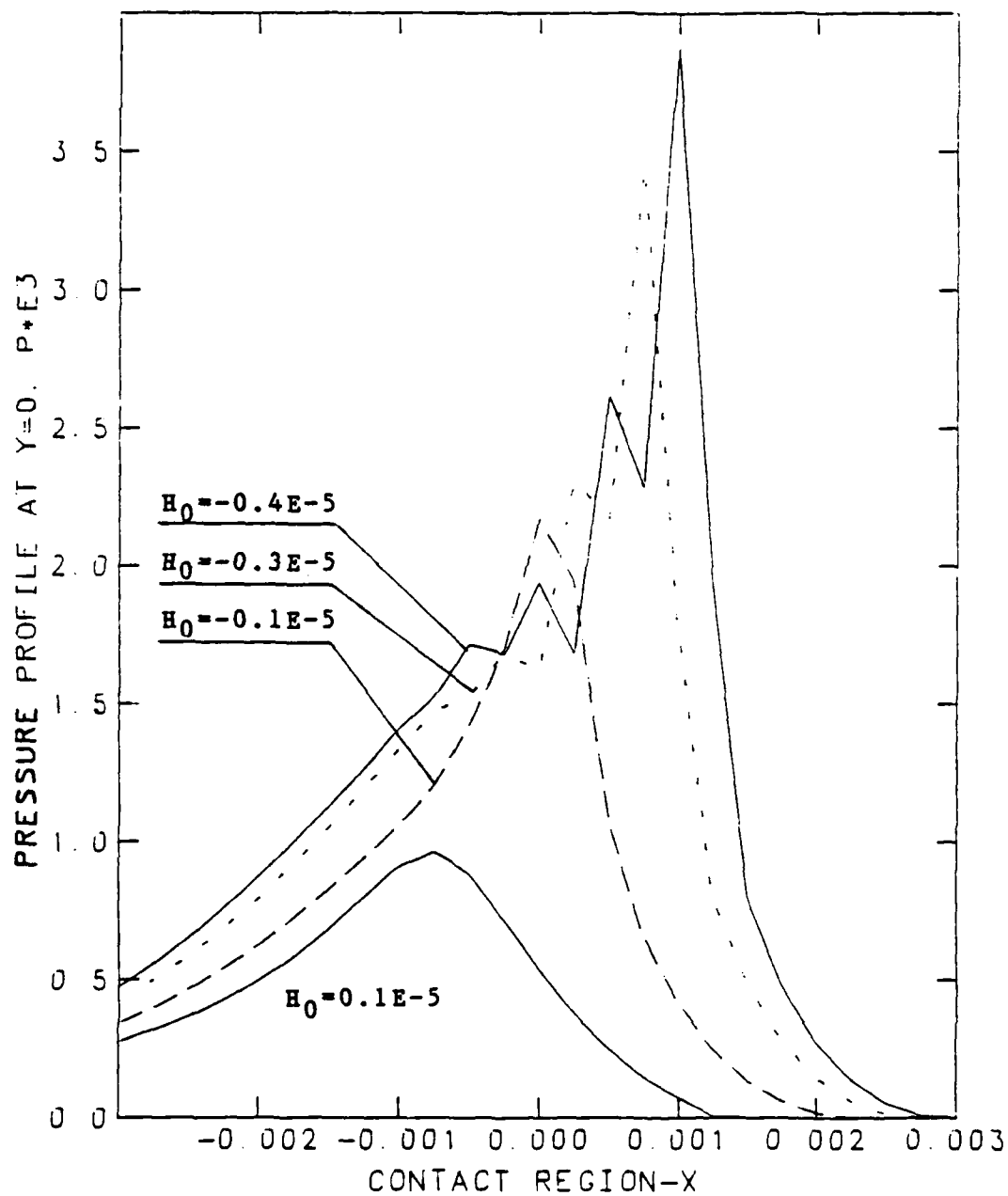
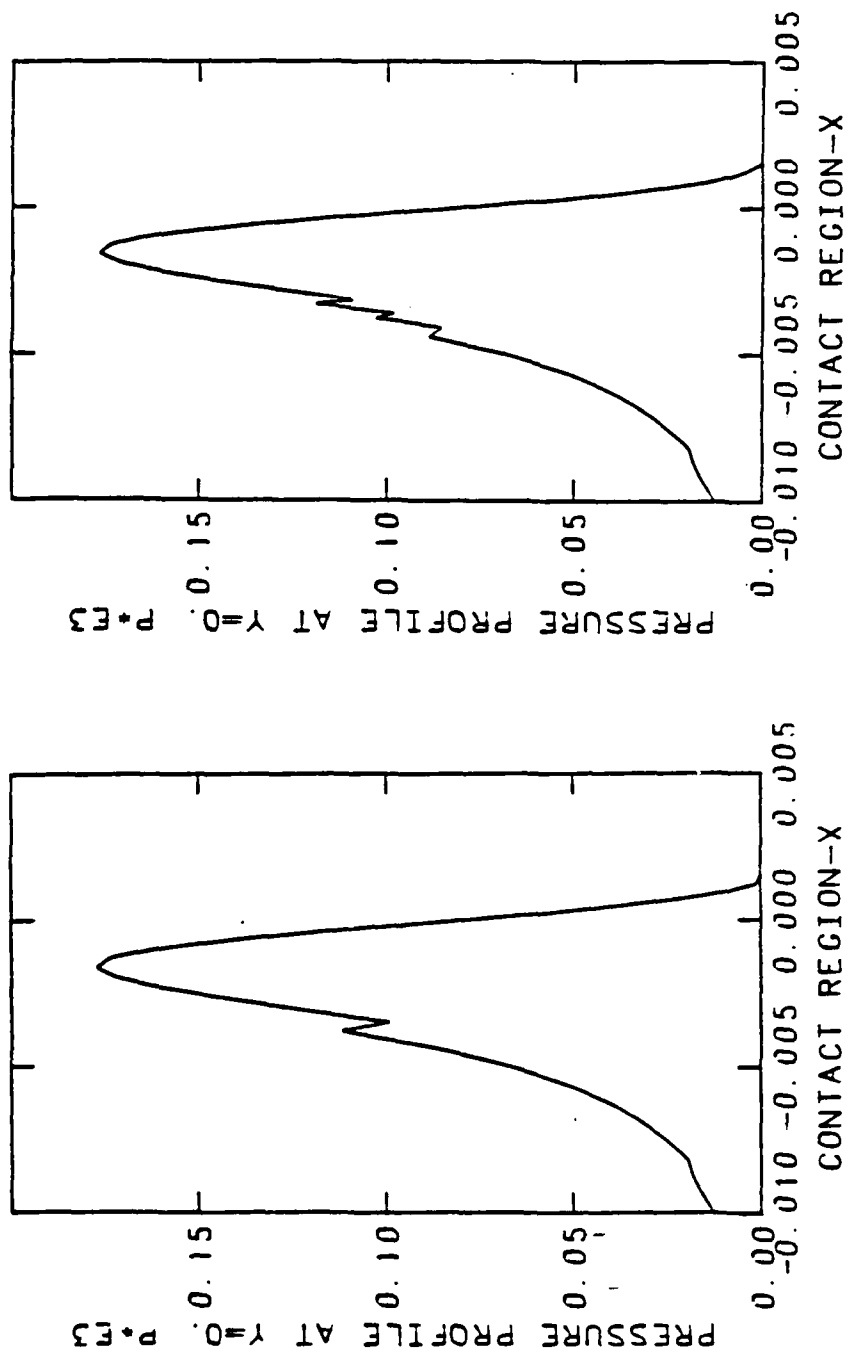


Fig. 7.1 Pressure Profiles of Point Contact under Heavy Load



(a) Third Refinement (242 elements) (b) Fourth Refinement (282 elements)
 Fig. 7.2 Wiggles Happened in Adaptive Refinement

REFERENCES TO CHAPTER 2

- [1] Antoniou, S.S., Cameron, A., and Gentle, C.R., "Friction-Speed Relation from Stick-Slip Data," *Wear*, 36, pp. 235-254, 1976.
- [2] Aronov, V., D'Souza, A.F., Kalpakjian, S., and Shareef, I., "Experimental Investigation on the Effect of System Rigidity on Wear and Friction-Induced Vibrations," *J. Lubr. Technol.*, 105, pp. 206-211, 1983.
- [3] Aronov, V., D'Souza, A. F., Kalpakjian, S., and Shareef, I., "Interactions Among Friction, Wear and System Stiffness — Part 1: Effect of Normal Load and System Stiffness; Part 2: Vibrations Induced by Dry Friction; Part 3: Wear Model," *J. Lubr. Technol.* 106, pp. 54-69, 1984.
- [4] Back, N., Burdekin, M., and Cowley, A., "Review of the Research on Fixed and Sliding Joints," *Proc. 13th International Machine Tool Design and Research Conference*, pp. 87-97, ed. by S. A. Tobias and F. Koenigsberger, MacMillan, London, 1973.
- [5] Bell, R., and Burdekin, M., "A Study of the Stick-Slip Motion of Machine Tool Feed Drives," *Proc. Inst. Mech. Engrs.*, 184(1), pp. 543-557, 1969-70.
- [6] Bell, R., and Burdekin, M., "An Investigation into the Steady-State Characteristics of Plain Slideways," *Proc. Instn. Mech. Engrs.*, 184-Pt 1(59), pp. 1075-1087, 1969-70.
- [7] Bhushan, B., "Effect of Strain Rate and Interface Temperature on the Static and Kinetic Frictions," *Seventh Leeds-Lyon Symposium*, Leeds, England, 1980.
- [8] Blok, H., "Fundamental Mechanical Aspects of Boundary Lubrication," *S.A.E. J.*, 46(2), pp. 54-68, 1940.
- [9] Bo, L.C. and Pavelescu, D., "The Friction-Speed Relation and its Influence on the Critical Velocity of Stick-Slip Motion," *Wear*, 82, pp. 277-289, 1982.
- [10] Bochet, M., "Nouvelles Recherches Expérimentales sur le Frottement de Glissement," *Ann. Mines*, xix, pp. 27-120, 1861.

- [11] Bowden, F.P., and Leben, L., "The Nature of Sliding and the Analysis of Friction," *Proc. Roy. Soc. London*, A169, pp. 371-391, 1939.
- [12] Bowden, F.P., and Tabor, D., "The Area of Contact Between Stationary and Between Moving Surfaces," *Proc. Roy. Soc. London*, A169, pp. 391-413, 1939.
- [13] Bowden, F.P., and Tabor, D., **The Friction and Lubrication of Solids**, Clarendon Press, Oxford, England, 1950.
- [14] Bowden, F.P., and Tabor, D., **The Friction and Lubrication of Solids, Part II**, Clarendon Press, Oxford, England, 1964.
- [15] Brockley, C.A., Cameron, R., and Potter, A.F., "Friction-induced Vibration," *J. Lubr. Technol.*, 89, pp. 101-108, 1967.
- [16] Brockley, C.A., and Davis, H.R., "The Time Dependence of Static Friction," *J. Lubr. Technol.*, 90, pp. 35-41, 1968.
- [17] Brockley, C.A., and Ko, P.L., "Quasi-Harmonic Friction-Induced Vibration," *J. Lubr. Technol., Trans. ASME*, Oct., pp. 550-556, 1970.
- [18] Budanov, B.V., Kudinov, V.A., and Tolstoi, D.M., "Interaction of Friction and Vibration," *Trenie i Iznos*, 1(1), pp. 79-89, 1980.
- [19] Burwell, J.T., and Rabinowicz, E., "The Nature of the Coefficient of Friction," *J. Appl. Phys.*, 24(2), pp. 136-139, 1953.
- [20] Chiou, Y.C., Kato, K., and Abé, H., "The Effect of Normal Damping in the Loading System on the Wear of Low Carbon Steel," *Wear*, 114, pp. 73-84, 1987.
- [21] Chiou, Y.C., Kato, K., and Kayaba, T., "Effect of Normal Stiffness in Loading System on Wear of Carbon Steel, Part 1, Severe-Mild Wear Transition," *J. Tribol.*, 107, p. 491, 1985.
- [22] Chiou, Y.C., and Kato, K., "Effect of Normal Stiffness in Loading System on Wear of Carbon Steel, Part 2, Dynamic Normal Load and Effective Sliding Distance," *J. Tribol.*, 108, p. 321, 1986.
- [23] Conti, P., "Sulla Resistenza di Attrito," *Atti R. Accad. Lincei*, 11, p. 16, 1875.

- [24] Coulomb, C.A., "Théorie des Machines Simples," *Mémoire de Mathématique et de Physique de l'Académie Royale*, pp. 161-342, 1785.
- [25] Coulomb, J., "L'Agitation Microsismique," *Encyclopedia of Physics*, XLVII, Ed. S. Flügge, Geophysics I, Group Ed. J. Bartels, Springer-Verlag, Berlin, Göttingen, Heidelberg, 1956.
- [26] Courtney-Pratt, J.S., and Eisner, E., "Effect of a Tangential Force on the Contact of Metallic Bodies," *Proc. Roy. Soc. Lond.*, A238, pp. 529-550, 1957.
- [27] Czichos, H., Becker, S., and Lexow, J., "Multi-laboratory Tribotesting: Results from the Versailles Advanced Materials and Standard Programme on Wear Test Methods," *Wear*, 114, pp. 109-130, 1987.
- [28] Dokos, S.J., "Sliding Friction Under Extreme Pressures — 1," *J. Appl. Mechanics*, 68, pp. A148-156, 1946.
- [29] Dowson, D., *History of Tribology*, Longman, London, 1978.
- [30] Earles, S.W.E., and Badi, M.N.M., "Oscillatory Instabilities Generated in a Double-pin and Disc Undamped System: a Mechanism of Disc-Brake Squeal," *Proc. Instn. Mech. Engrs.*, 198C(4), pp. 43-50, 1984.
- [31] Earles, S.W.E., and Lee, C.K., "Instabilities Arising from the Frictional Interaction of a Pin-Disk System Resulting in Noise Generation," *J. of Engineering for Industry, Transactions of the ASME*, pp. 81-86, Feb. 1976.
- [32] Elder, J.A., Jr., and Eiss, N.S., Jr., "A Study of the Effect of Normal Stiffness on Kinetic Friction Forces Between Two Bodies in Sliding Contact," *ASLE Transactions*, 12, pp. 234-241, 1969.
- [33] Euler, L., "Sur le Frottement des Corps Solides," *Histoire de l'Académie Royale à Berlin*, iv (1748), pp. 313, 1750.
- [34] Galton, D., "The Action of Brakes," *Engineering*, 23, pp. 153-154; 25, pp. 469-472, 1878.
- [35] Gu, J.-C., Rice, J.R., Ruina, A.L., and Tse, S.T., "Slip Motion and Stability of a Single Degree of Freedom Elastic System with Rate and State Dependent Friction," submitted to *J. Mech. Phys. Solids*, 1983.

- [36] Guckenheimer, J., and Holmes, P., **Nonlinear Oscillations, Dynamical System, and Bifurcation Theory**, Springer, Berlin, 1983.
- [37] Heyman, F., Rabinowicz, E., and Rightmire, B.G., *Rev. Sci. Instrum.*, 26, p. 56, 1955.
- [38] Hirn, G., "Sur les Principaux Phénomènes qui Présentent les Frottements Médiats," *Bull. Soc. Ind. Mulhouse*, 26, pp. 188-277, 1854.
- [39] Hunt, J.B., Torbe, I. and Spencer, G.C., "The Phase-Plane Analysis of Sliding Motion," *Wear*, 8, pp. 455-465, 1965.
- [40] Hunt, K.H., and Crossley, F.R.E., "Coefficient of Restitution Interpreted as Damping in Vibroimpact," *J. Appl. Mech.*, pp. 440-445, June 1975.
- [41] Huseyn, K., **Vibrations and Stability of Multiple Parameter Systems**, Noordhoff International Publishing, Alphen Aan Den Rijn, The Netherlands, 1978.
- [42] Ishlinski, A.Y. and Kraghelsky, I.V., *J. Techn. Phys. (Russian)*, 14, p. 276, 1944.
- [43] Jarvis, R.P., and Mills, B., "Vibrations Induced by Dry Friction," *Proc. Instn. Mech. Engrs.*, 178-Pt 1(32), pp. 847-866, 1963-64.
- [44] Jenkin, F., and Ewing, J.A., "On Friction Between Surfaces Moving at Low Speeds," *Phil. Trans. R. Soc. Lond.*, 167, pp. 509-528, 1877.
- [45] Johannes, V.I., Green, M.A., and Brockley, C.A., "The Role of the Rate of Application of the Tangential Force in Determining the Static Friction Coefficient," *Wear*, 24, pp. 384-385, 1973.
- [46] Kaidanovski, N.L., and Khaikin, S.E., *J. Techn. Phys. (Russian)*, 3, p. 1, 1933.
- [47] Kato, S., and Matsubayashi, T., "On the Dynamic Behavior of Machine-Tool Slideway - 1st Report," *Bull. J.S.M.E.*, 13(55), pp. 170-179, 1970.
- [48] Kato, K., Iwabuchi, A., and Kayaba, T., "The Effects of Friction-induced Vibration on Friction and Wear," *Wear*, 80, pp. 307-320, 1982.
- [49] Kimball, A.L., "Sliding Friction on an Inclined Plane," *Am. J. Sci. Arts*, xiii (Silliman's Journal), p. 151; "A New Investigation on the Laws of Friction," *Am. J. Sci. Arts*, xiii (Silliman's Journal), p. 353, 1877.

- [50] Kragelskii, I.V., **Friction and Wear**, Butterworths, Washington, 1965.
- [51] Kragelskii, I.V., Dobychin, M.N., and Kombalov, V.S., **Friction and Wear: Calculation Methods**, Pergamon Press, 1982.
- [52] Kudinov, V.A., "An Investigation of the Vibrations of Machine Tools," Mashgiz (State Mech. Eng. Press), p. 251, 1958.
- [53] Leipholz, H.H.E., **Stability Theory**, Academic Press, Inc., New York and London, 1970.
- [54] Madakson, P.B., "The Frictional Behavior of Materials," *Wear*, 87, pp. 191-206, 1983.
- [55] Morin, A.J., *Mém. Savans Etrang.*, (Paris), iv, pp. 1-128; iv, pp. 591-696; vi, pp. 641-785; *Ann. Min.*, iv, pp. 271-321; vi, pp. 73-96; x, pp. 27-56, 1833-36.
- [56] Oden, J.T., and Martins, J.A.C., "Models and Computational Methods for Dynamic Friction Phenomena," *Computer Methods in Appl. Mechs. and Engng.*, 52, pp. 527-634, 1985.
- [57] Poirée, *Mém. Soc. Ing. Civ. Fr.*, 1852.
- [58] Rabinowicz, E., "The Nature of the Static and Kinetic Coefficients of Friction," *J. Appl. Physics*, 22(11), pp. 1373-1379, 1951.
- [59] Rabinowicz, E., "The Intrinsic Variables Affecting the Stick-Slip Process," *Proc. Phys. Soc.*, 71, pp. 668-675, 1958.
- [60] Rabinowicz, E., **Friction and Wear of Materials**, John Wiley and Sons, New York, 1965.
- [61] Rennie, G., "Experiments on the Friction and Abrasion of the Surfaces of Solids," *Phil. Trans. R. Soc. Lond.*, 34, Pt I, pp. 143-170, 1829.
- [62] Rice, S.L., Hans, N., and Steven, F.W., "The Role of Specimen Stiffness in Sliding and Impact Wear," *Wear*, 77, pp. 13-28, 1982.
- [63] Rice, J.R., and Ruina, A.L., "Stability of Steady Frictional Slipping," Paper No. 83-APM-16, *ASME Applied Mechanics, Bioengineering, and Fluids Engineering Conference*, Houston, Tx, 1983.

- [64] Richardson, R.S.H., and Nolle, H., "Surface Friction Under Time-dependent Loads," *Wear*, 37, pp. 87-101, 1976.
- [65] Ruina, A.L., "Friction Laws and Instabilities: A Quasistatic Analysis of Some Dry Frictional Behavior," Ph.D. Thesis, Brown University, Providence, RI, 1980.
- [66] Ruina, A.L., "Slip Instability and State Variable Friction Laws," *J. Geophys. Res.*, 1983.
- [67] Shobert, E.I., "Carbon Brush Friction and Chatter," *Trans. Amer. Inst. Elect. Engrs.*, 76, Part III, p. 268, 1957.
- [68] Simkins, T.E., "The Mutuality of Static and Kinetic Friction," *Lubr. Engrg.*, 23, pp. 26-31, 1967.
- [69] Soda, N., Kimura, Y., and Tanaka, A., "Wear of Some f.c.c. Metals During Unlubricated Sliding. Part I: Effects of Load, Velocity and Atmosphere Pressure on Wear," *Wear*, 33, pp. 1-16, 1975.
- [70] Soom, A., and Kim, C., "Interactions Between Dynamic Normal and Friction Forces During Unlubricated Sliding," *J. Lubr. Technol.*, 105, pp. 221-229, 1983.
- [71] Soom, A., and Kim, C., "Roughness-induced Dynamic Loading at Dry and Boundary-lubricated Sliding Contacts," *J. Lubr. Technol.*, 105, pp. 514-517, 1983.
- [72] Spurr, R.T., "A Theory of Brake Squeal," *Proc. Instn. Mech. Engrs. (A.D.)*, 1, pp. 33-52, 1961-62.
- [73] Thornley, R.H., Connolly, R., Barash, M.M., and Koenigsberger, F., "The Effect of Surface Topography Upon the Static Stiffness of Machine Tool Joints," *Int. J. Mach. Tool Des. Res.*, 5, pp. 57-74, 1965.
- [74] Tolstoi, D.M., "Significance of the Normal Degree of Freedom and Natural Normal Vibrations in Contact Friction," *Wear*, 10, pp. 199-213, 1967.
- [75] Tolstoi, D.M., Borisova, G.A., and Grigorova, S.R., "Role of Intrinsic Contact Oscillations in Normal Direction During Friction," *Nature of the Friction of Solids*, Nauka i Tekhnika, p. 116, Minsk, 1971.

- [76] Tse, S.T., and Rice, J.R., "Crustal Earthquake Instability in Relation to the Depth Variation of Frictional Slip Properties," *21st Annual Meeting, Society of Engineering Science, Inc.*, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1984.
- [77] Tudor, A., and Bo, L.C., "The Squeeze Film Under Boundary Lubrication Conditions and its Effect on the Vertical Displacement of Sliding Bodies," *Wear*, 80, pp. 115-119, 1982.
- [78] Vince, S., "On the Motion of Bodies Affected by Friction," *Phil. Trans. R. Soc. Lond.*, 75, pt. I, pp. 165-189, 1785.
- [79] Ziegler, H., "Die Stabilitätskriterien der Elastomechanik," *Ingenieur-Archiv.*, 20, pp. 49-56, 1952.
- [80] Ziegler, H., *Principles of Structural Stability*, Blaisdell Publ. Co., 1968.
- [81] Sampson, J.B., Morgan, F., Reed, D.W., and Muskat, M., "Friction Behavior During the Slip Portion of the Stick-Slip Process," *J. Appl. Phys.*, 14, pp. 689-700, 1943.
- [82] Kudinov, V.A., and Lisitsyn, N.M., "Factors Affecting Uniform Movement of Tables and Tool Slides in Mixed Friction Conditions," *Mach. Tool*, 33(2), pp. 2-6, 1962.
- [83] Klarbring, A., *Contact Problems in Linear Elasticity*, Linköping Studies in Science and Technology, Dissertations, No. 133, 1985.
- [84] Ruina, A.L., *Wear*, 113, pp. 83, 1986.
- [85] Boyer and Cameron, "A Computer Sliding Model of Rough Surfaces," *Leeds-Lyon Symposium*.
- [86] Watari, A., and Sugimoto, T., "Vibrations Caused by Dry Friction," *Bulletin of J.S.M.E.*, 7, pp. 40-52, 1964.
- [87] Yokoi, M., and Nakai, M., "A Fundamental Study on Frictional Noise," *Bull. of J.S.M.E.*, Vol. 22, No. 173, pp. 1665-1671, 1979.
- [88] Martins, J. A. C., *Dynamic Frictional Contact Problems Involving Metallic Bodies*, Ph.D Dissertation, The University of Texas at Austin, U.S.A., 1986.

REFERENCES TO CHAPTER 3

- [1] AMONTONS G. [1699], De la Résistance Causé dans les Machines, *Memoires de l'Academie Royale A*, pp. 275-282.
- [2] ANAND A., SOOM A., [1984], Roughness-induced Transient Loading at a Sliding contact During Start-up, *Journal of Tribology*, **106**, pp. 69-53.
- [3] ANDREW C., COCKBURN J.A. and WARING A.E. [1968], Metal Surfaces in Contact Under Normal Forces: Some Dynamic Stiffness and Damping Characteristics, *Proc. Instn. Mech. Engrs.*, **182**(3K), pp. 92-100.
- [4] ANTONIOU, S. S., CAMERON A., and GENTLE, D. R., [1976], Friction-Speed Relation from Stick-Slip Data, *Wear*, **36**, pp. 235-254.
- [5] ARANOV V., D'SOUZA A.F., KALPAKIJAN S. and SHAREEF I. [1983], Experimental Investigation on the Effect of System Rigidity on Wear and Friction-Induced Vibrations, *Journ. Lubr. Technol.*, **105**, pp. 206-211.
- [6] ARANOV V., D'SOUZA A.F., KALPAKIJAN S. and SHAREEF I. [1984], Interaction Among Friction, Wear and System Stiffness - Part1: Effect of Normal Load and System Stiffness; Part 2: Vibrations Induced by Dry Friction; Part 3: Wear Model, *Journ. Lubr. Technol.*, **106**, pp. 54-69.
- [7] BACK N., BURDEKIN M. and COWLEY A. [1973], Review of the Research on Fixed and Sliding Joints, *Proc. 13th Internat. Machine and Tool Design and Research Conference*, pp. 87-97, ed. by S.A. Tobias and F. Koenigsberger, McMillan, London.
- [8] BAY N., WANHEIM T. [1976], Real area of contact and Friction Stress at High Pressure Sliding Contact, *Wear*, **38**, pp. 201-209.
- [9] BECKER E., TWORZYDLO, W. W. [1988], Numerical Analysis of Dynamic Friction, TICOM report, The University of Texas at Austin, in preparation.
- [10] BELL R. and BURDEKIN M. [1969-70] A Study of the Stick-Slip Motion of Machine Tool Feed Drives, *Proc. Inst. Mech. Engrs.*, **184**(1), pp. 543-557.
- [11] BHUSHAN B., Effect of Strain Rate and Interface Temperature on the Static and Kinetic Frictions [1980], *Seventh Leeds-Lyon Symposium*, Leews, England.
- [12] BOWDEN F.P. and TABOR D. [1950], *The Friction and Lubrication of Solids*, Clarendon Press, Oxford.

- [13] BOWDEN F.P. and TABOR D. [1964], *The Friction and Lubrication of Solids - Part II*, Clarendon Press, Oxford.
- [14] BÖHM J., [1987], A Comparison of Different Contact Algorithms with Applications, *Computers and Structures*, **26**, pp. 207-221.
- [15] BROCKLEY C.A. and DAVIS H.R. [1968] , The Time Dependence of Static Friction, *Journ. Lubr. Technol.*, **90**, pp. 35-41.
- [16] BRONIEC Z. and LENKIEWICZ W. [1980], Static Friction Process Under Dynamic Loads and Vibration, *Wear*, **80**, pp. 261-271.
- [17] BUCKLEY D. [1977], The Metal-to-Metal Interface and Its Effect on Adhesion and Friction, *Journ. of Colloid and Interface Science*, **58**, pp. 36-53.
- [18] BURWELL J.T. and RABINOWICZ E. [1953], The Nature of the Coefficient of Friction, *Journ. Appl. Phys.*, **24**, 2, pp. 136-139.
- [19] COULOMB C.A. [1785], Théorie de Machines Simples, *Mémoire de Mathématique et de Physique de l'Academie Royale*, pp. 161-342.
- [20] CHANG W.R., ETSION I. and BOGY D.B. [1986] Static Friction Coefficient Model for Metallic Rough Surfaces, *University of California at Berkeley, Rep. No UCB-MICRO 85-126* , **3**, pp. 1-28.
- [21] CHOIU, Y. C., KATO, L., ABE, H., [1987], The Effect of Normal Damping in the Loading System on the Wear of Low Carbon Steel, *Wear*, **114**, pp. 73-84.
- [22] CONNOLLY R., SCHOFIELD R.E. and THORNLEY R.H. [1967], The Approach of Machined Surfaces with Particular Reference to their Hardness, *Advances in Machine Tool Design and Research*, Proc. 8th Intern. M.T.D.R. Conf., Univ. of Manchester, Sept. 1967, Part 2, pp. 759-775, ed. by S.A. Tobias and F. Koenigsberger, Pergamon Press, Oxford.
- [23] DONEA J., GIULIANI S. and HALLEUX J.P., [1982], An Arbitrary Lagrangian-Eulerian Finite Element Method for Transient Dynamic Fluid-Structure Interactions , *Comp. Meth. Appl. Mech. Engng.*, **33**, pp. 689-723.
- [24] DUVAUT, G., LIONS, J. L., [1976], Inequalities in Mechanics and Physics, *Springer-Verlag, Berlin, Heidelberg, New York*.
- [25] COURTNEY-PRATT J.S. and EISNER E. [1957], The Effect of Tangential Force on the Contact of Metallic Bodies, *Proc. Royal Soc. Lond.*, A238, pp. 529-550.

- [26] GODFREY D. [1967], Vibration Reduces Metal to Metal Contact and Causes an Apparent Reduction in Friction, *ASLE Transactions*, **10**, pp. 183-192.
- [27] GRIGOROVA S.R. and TOLSTOI D.M. [1966], *Dokl. Akad. Nauk SSSR*, **167**, 3, pp. 562-564.
- [28] HUNT K.H. and CROSSLEY F.R.E. [1975], Coefficient of Restitution Interpreted as Damping in Vibroimpact, *Journ. Appl. Mech.*, June, 1975 ??
- [29] JOHANNES V.I., GREEN M.A. and BROCKLEY C.A. [1973], The Role of the Rate of Application of the Tangential Force in Determining the Static Friction Coefficient, *Wear*, **24**, pp. 384-385. ?? pages
- [30] KATO S. and MATSUBAYACHI T. [1970], On the Dynamic Behaviour of Machine-Tool Slideway – 1st Report, *Bull. J.S.S.M.E.*, **13**, **55**, pp. 170-179.
- [31] KATO S., YAMAGUCHI K. and MATSUBAYACHI T. [1970], On the Dynamic Behaviour of Machine-Tool Slideway – 2ND Report, *Bull. J.S.S.M.E.*, **13**, **55**, pp. 180-188.
- [32] KATO S., SATO N. and MATSUBAYACHI T. [1972], Some Considerations on Characteristics of Static Friction of Machine Tool Slideway, *Journ. Lubr. Technol.*, **94**, pp. 234-247.
- [33] KENNEDY J.M. and BELYTSCHKO T.B., [1981], Theory and Application of a Finite Element Method for Arbitrary Lagrangian-Eulerian Fluids and Structures, *Nuclear Engng. and Design*, **68**, pp 129-146.
- [34] KIKUCHI, and N., ODEN, J. T., [1988], Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods, *SIAM, Philadelphia*.
- [35] KO P.L. and BROCKLEY C.A. [1970] , The Measurment of Friction and Friction-Induced Vibration, *Journ. Lubric. Technol.*, **92**, pp. 543-548.
- [36] LANCHUN, H., MAKAYA, B., MIRGAUX, A., SAINT JEAN PAULIN, J., and KRÖNER, E., [1985], A Mathematical Study to Obtain Quantitative Effects of Roughness in Technical Problems, *Wear*, **109**, pp. 99-111.
- [37] LENKIEWICZ W.[1969], The Sliding Friction Process – Effect of External Vibrations, *Wear*, **13**, pp. 99-108.
- [38] LIFSHITZ J.M. and KOLSKY K. [1964] , Some Experiments on Anelastic Rebound, *Journ. Mech. Phys. of Solids*, **12**, pp. 35-43.

- [39] NOLLE H. and RICHARDSON R.S. [1974], Static Friction Coefficient for Mechanical and Structural Joints, *Wear*, 28, pp. 1-13.
- [40] ODEN J.T. and MARTINS J.A.C. [1985], Models and Computational Methods for Dynamic Friction Phenomena, *Comp. Meth. Appl. Mech. and Engng.*.
- [41] PANAGIATOPOULOS, P. D., [1975], A Nonlinear Programming Approach to the Unilateral Contact and Friction-Boundary-Value Problems in the Theory of Elasticity, *Ing.Archiv*, 44, pp. 421-532.
- [42] POHLMAN, R., and LEHFELDT, E., [1965], Die Einfluss von Ultraschall-Schwingungen auf metallische Reibungsvorgänge, (Lab. für Ultrasch., T. H. Aachen), *Rept 5 Congr. Intern d'Acoustique*, Liège.
- [43] RABINOWICZ E. [1951], The Nature of the Static and Kinetic Coefficients of Friction, *Journ. Appl. Physics*, 11, 22, pp. 1373-1379.
- [44] RABINOWICZ E. [1957], The Intrinsic Variables Affecting the Stick-Slip Process, *Proc. Phys. Soc.*, 71, pp. 668-675.
- [45] RABINOWICZ E. [1959], A Study of the Stick-Slip Process, *Symp. on Friction and Wear*, Detroit, 1957, pp. 149-161, ed. by R. Davies, Elsevier, New York.
- [46] RICHARDSON R.S.H. and NOLLE H. [1976], Surface Friction Under Time-Dependent Loads, *Wear*, 37, pp. 87-101.
- [47] RUINA A.L. [1985] Constitutive Relations for Frictional Slip, *Mechanics of Geomaterials*, ed. by Z. Bazant, John Wiley and Sons, pp. 163-187.
- [48] SAMPSON J.B., MORGAN F., REED D.W. and MUSKAT M. [1943] Friction Behaviour During the Slip Portion of the Stick-Slip Process, *Journ. Appl. Physics*, 14, pp. 689-700.
- [49] SCHREURS P.J.G., VELDPAUS F.E. and BREKELMANS W.A.M., [1986], Simulation of Forming Process Using the Arbitrary Eulerian-Lagrangian Formulation, *Comp. Meth. Appl. Mech. Engng.*, 58, pp. 19-36.
- [50] SOOM, A., CHEN, J. W., [1986], Simulation of Random Surface Roughness-Induced Contact vibrations at Hertzian Contact During Steady Sliding, *Trans. ASME*, pp. 123-127.
- [51] SOOM A. and KIM C. [1983a], Interactions Between Dynamic Normal and Frictional Forces During Unlubricated Sliding, *Journ. Lubr. Technol.*, 105, pp. 221-229.

- [52] SOOM A. and KIM C. [1983b], Roughness-Induced Dynamic Loading at Dry and Boundary-Lubricated Sliding Contacts, *Journ. Lubr. Technol.*, 105, pp. 514-517.
- [53] TABOR D. [1981] Friction - the Present State of Our Understanding, *Journ. Lubr. Technol.*, 103, pp. 169-179.
- [54] R.H., CONOLLY R., BARASCH M.M. and KOWNIGSBERGER F., [1965], The effect of Surface Topography Upon the Static Stiffness of MACHINE Tool Joints, *Int. Journ. Mach. Des. Tool Res.*, 5, pp. 55-74.
- [55] TOLSTOI D.M. [1967], Significance of the Normal Degree of Freedom and Natural Normal Vibrations in Contact Friction, *Wear*, 10, pp. 199-213.
- [56] TOLSTOI D.M., BORISOVA G.A. and GRIGOROVA S.R. [1971], Role of Intrinsic Contact Oscillations in Normal Direction During Friction, *Nature of the Friction of Solids*, Nauka i Tiechnika, p.116, Minsk.
- [57] TWORZYDLO, W. W., BECKER, E., [1988], Influence of Forced Vibrations on the Static Coefficient of Friction-Numerical Analysis, *prepared for publication in Wear*.
- [58] WEIC, W. L. [1962], The Investigation of the Static Friction and the Direction of Sliding at Low-Frequency External Vibrations, *New in Friction Theory*, Lzd. Akad, Nauk SSSR, Moscow, pp. 60-82.
- [59] WILSON [1952] , Ph. D. Thesis, *University of Cambridge*.

REFERENCES TO CHAPTER 4

- [1] Amontons, G., "De la resistance causee dans les machines", *Memoires de l'Academie Royale*, A, p. 275-282, 1699.
- [2] Coulomb, C.A., "Theorie des machines simples", *Memoire de Mathematique et de Physique de l'Academie Royale*, p. 161-342, 1785.
- [3] Hertz, H., "Uber die Beruhrung fester elastische Korper und uber die Harte", *Verhandlungen des Vereins zur Beforderung des Gewerbefleisses Leipzig*, Nov. 1882.
- [4] Tabor, D., "Friction - The Present State of Our Understanding", *ASME Journal of Lubrication Technology*, Apr, 103, p. 169-179, 1981.
- [5] Abbott, E.J. and Firestone, F. A., "Specifying Surface Quality", *Mechanical Engineering*, Vol. 55, p. 569, 1933.
- [6] Bowden, F. P. and Tabor, D., "The Friction and Lubrication of Solids", (Claredon Press, Oxford, 1950.)
- [7] Bowden, F.P. and Tabor, D., "The Friction and Lubrication of Solids, Part II", (Claredon Press, Oxford, 1964.)
- [8] Greenwood, J.A., and Williamson, J.B.P., "Contact of Nominally Flat Surfaces", *Proc. Roy. Soc. (London)*, Series A295, p. 300-319, 1966.
- [9] Mindlin, R.D., "Compliance of Elastic Bodies in Contact", *Journal of Applied Mechanics*, 16, p. 259-268, 1949.
- [10] Greenwood, J.A. and Tripp, J. H., "The Elastic Contact of Rough Spheres", *Journal of Applied Mechanics*, 34, p.134-159, 1967.
- [11] Johnson, K.L., Kendall, K., and Roberts, A. D., "Surface Energy and the Contact of Elastic Solids", *Proc. Roy. Soc. (London)*, A234, p. 301-313, 1971.
- [12] Derjaguin, B. V., Muller, V.M., and Toporov, Y. P., "Effect of Contact Deformations on the Adhesion of Particles", *Journal of Colloid and Interface Science*, 53, p.314-326, 1975.
- [13] Muller, V.M., Yushchenko, V.S., and Derjaguin, B. V., "On the Influence of Molecular Forces on the Deformation of an Elastic Sphere and Its Sticking to a Rigid Plane", *Journal of Colloid and Interface Science*, 77, p. 91-101, 1980.
- [14] Muller, V.M., Derjaguin, B. V., and Toporov, Y.P., "On Two Methods of Calculation of the Force of Sticking of an Elastic Sphere to a Rigid Plane", *Colloids and Surfaces*, 7, p. 251, 1983.

- [15] Briscoe, B.J., Scruton, B., and Willis, F. R., "The Shear Strength of Thin Lubricant Films", *Proc. Roy. Soc. (London)*, A333, p. 99-114, 1973.
- [16] Briscoe, B.J., and Tabor, D., "The Effect of Pressure on the Frictional Properties of Polymers", *Wear*, 34, p. 29 - 38, 1975.
- [17] Francis, H.A., "Phenomenological Analysis of Plastic Spherical Indentation", *ASME Journal of Engineering Material Technology*, 98, p. 272 - 281, 1976.
- [18] Francis, H.A., "Application of Spherical Indentation Mechanics to Reversible and Irreversible Contact between Rough Surfaces", *Wear*, 45, p. 261 - 269, 1977.
- [19] Chang, W.R., Etsion, I., and Bogy, D.B., "An Elastic-Plastic Model for the Contact of Rough Surfaces", *University of California Berkeley Report*, No. UCB-MICRO 85-126 #1, March, 1986.
- [20] Chang, W.R., Etsion, I., and Bogy, D. B., "Adhesion Model for Metallic Rough Surfaces", *University of California Berkeley Report*, No. UCB-MICRO 85-126 #2, March, 1986.
- [21] Chang, W.R., Etsion, I., and Bogy, D. B., "Static Friction Coefficient Model for Metallic Rough Surface", *University of California Berkeley Report*, No. UCB-MICRO 85-126 #3, March, 1986.
- [22] Whitehouse, D. J. and Archard, J.F., "The Properties of Random Surfaces of Significance in their Contact", *Proc. Roy. Soc. (London)*, A316, p. 97-121, 1970.
- [23] Nayak, P.R., "Random Process Model of Rough Surfaces", *ASME Journal of Lubrication Technology*, 93, p. 398-407, 1971.
- [24] Bush, A.W., Gibson, R. D., and Thomas, T.R., "The Elastic Contact of a Rough Surface", *Wear*, 35, p. 87 - 111, 1975.
- [25] Hisakado, T., "Effect of Surface Roughness on Contact between Solid Surfaces", *Wear*, 28, p. 217 - 234, 1974.
- [26] Hisakado, T., and Tsukizoe, T., "Effects of Distribution of Surface Slopes and Flow Pressures of Contact Asperities on Contact between Solid Surfaces", *Wear*, 30, p. 213 - 227, 1974.
- [27] Halling, J. and Nuri, K.A., in dePater, A.D. and Kalker, J.J. (Eds.) "The Mechanics of the Contact between Deformable Bodies", *Delft University Press*, Delft, p. 330-341, 1975.
- [28] Tolstoi, D.M., "The Significance of the Normal Degree of Freedom and Natural Normal Vibrations in Contact Friction", *Wear*, 10, p. 199 - 213, 1967.
- [29] Oden, J.T. and Martins, J.A.C., "Models and Computational Methods for Dynamic Friction Phenomena", *Computer Methods in Applied Mechanics and Engineering*, 52, p.527 - 634, 1985.
- [30] Signorini, A., "Sopra alcune questioni di elastostatica", *Atti della Societa Italiana per*

il Professo delle Scienze, 1933.

- [31] Signorini, A., "Questioni de elasticita nonlinearizzata e semi linearizzata", Rend. de Malematica, Rome, 1959.
- [32] Fichera, G., "Problemi elastostatici con vincoli unilaterali II: Problema di signorini con ambigue condizioni al contorno", Mem. Accad. Naz. Lincei, S. VIII, Vol. VII, Sez. I, 5, p. 91-140, 1964.
- [33] Duvaut, G. and Lions, J.L., "Les inequations en mecanique et en physique", Dunod, Paris, 1972.
- [34] Necas, J., Jarusek, and Haslinger, J., "On the Solution of the Variational Inequality to the Signorini Problem with Small Friction", Boll. UMI, 17-B, p. 796 - 811, 1986.
- [35] Duvaut, G., "Problems Mathematiques de la Mecanique-equilibre dun solide elastique avec contact unilateral et frottement de Coulomb", C.R. Acad. Sci., Paris, A, 296, p. 263 - 265, 1980.
- [36] Oden, J.T. and Pires, E., "Numerical Analysis of Certain Contact Problems in Elasticity with Non-Classical Friction Laws", Compt. & Struct., 16, 1983, Adv. and Trends in Structures and Solid Mechanics, Washington, D.C., October 4 - 7, p. 481 - 485, 1982.
- [37] Oden, J.T. and Pires, E., "Nonlocal and Nonlinear Friction Laws and Variational Principles for Contact Problems in Elasticity", Trans. ASME Ser. E, Journal Applied Mechanics, 50, p. 67 - 76, 1983.
- [38] Demkowicz, L. and Oden, J.T., "On some existence and uniqueness results in contact problems with nonlocal friction", Journal of Nonlinear Analysis, 1981.
- [39] Martins, J.A.C. and Oden, J.T., "Existence and Uniqueness Results for Dynamic Contact Problems with Nonlinear Normal and Friction Interface Laws, Journal of Nonlinear Analysis, 1987 (to appear).
- [40] Martins, J.A.C., "Dynamic Frictional Contact Problems Involving Metallic Bodies", Ph.D. Dissertation, The University of Texas, Austin, Texas, 1985.
- [41] Rabier, P., Martins, J.A.C., Oden, J.T., and Campos, L., "Existence and Local Uniqueness of Solutions to Contact Problems in Elasticity with Nonlinear Friction Laws", International Journal of Engineering Science, Vol. 24, No. 11, P. 1755 - 1768, 1986.
- [42] Martins, J.A.C. and Oden, J.T., "A Simple Method for Dynamic Friction Effects on Metallic Surfaces, Constitutive Laws for Engineering Materials - Theory and Applications, University of Arizona, January, 1987.
- [43] Kikuchi, K. and Oden, J.T., Contact Problems in Elasticity, SIAM Publications, Philadelphia, 1987.
- [44] Bodner, S.R., and Partom, Y., "A Large Deformation Elastic-Viscoplastic Analysis of a Thick Walled Spherical Shell", Journal of Applied Mechanics, 39, p. 751, 1972.

- allian, T.E., *Wear*, 21, p. 49, 1972.
- ashley, M.D., "Further Consideration of the DMT Model for Elastic Contact", *Colloids and Surfaces*, 12, p. 69 - 77, 1984.
- ashley, M.D., Pethica, J.B., and Tabor, D., "Adhesion and Micromechanical Properties of Metal Surfaces", *Wear*, 100, p. 7 - 31, 1984.
- ashley, M.D. and Pethica, J.B., "The Role of Surface Forces in Metal - Metal Contacts", *Journal Vac. Sci. Tech.*, A3, p. 757 - 761, 1985.
- Rabinowicz, E., "The Determination between Solids in Vacuum; II. Deformation and Interfacial Energy", *Journal of Physics*, V. 11, P. 39 - 54, 1978.
- Ferrante, J., Smith, J. and Rose, J.H., "Metallic Adhesion and Bonding", *Tribology in the 80's*, NASA CP2300, p. 143 - 163, 1983.
- Fuller, K.N.G. and Tabor, D., "The Effect of Surface Roughness on the Adhesion of Elastic Solids," *Proc. Roy. Soc. (London)*, A345, p. 327 - 342, 1975.
- McFarlane, J.S. and Tabor, D., "Adhesion of Solids and the Effect of Surface Films", *Proc. Roy. Soc. (London)*, A202, p. 224 - 243, 1950.
- McFarlane, J.S. and Tabor, D., "Relation between Friction and Adhesion", *Proc. Roy. Soc. (London)*, A202, p. 244 - 253, 1950.
- Bass, J.M. and Oden, J.T., "Adaptive Finite Element Methods for a Class of Evolution Problems in Viscoplasticity", *International Journal of Engineering Science*, (in press.)
- 1) Bodner, S.R. and Stouffer, D.C., "Constitutive Model for the Deformation Induced Anisotropic Plastic Flow of Metals", *International Journal of Engineering Science*, 17, p. 757-764, 1979.
- 2) Boyd, J. and Robertson, B.P., "The Friction Properties of Various Lubricants at High Pressures", *Trans. ASME*, 67, pp. 51-59, 1945.
- 7) Towle, L.C., "Shear Strength and Polymer Friction", *Am. Chem. Soc. Polymer Science and Technology*, Vol. 5A, pp. 179-189, 1974.
- 8) Villaggio, P., "An Elastic Theory of Coulomb Friction", *Archive for Rational Mechanics and Analysis*, Vol. 70, P. 135-143, 1979.

REFERENCES TO CHAPTER 5

- [1] Duvaut, G. and Lions, J. L., *Inequalities in Mechanics and Physics*, Springer-Verlag, Berlin, Heidelberg, New York, 1976.
- [2] Lötstedt, P., "Coulomb Friction in Two-Dimensional Rigid Body Systems", *ZAMM*, 61 pp. 605-615, 1984.
- [3] Jean, M., and Pratt, E., "A System of Rigid Bodies with Dry Friction", *International Journal of Engineering Sciences*, Vol. 23, No.5, pp. 497-513, 1985.
- [4] Schatzmann, M., "A Class of Nonlinear Differential Equations of Second Order in Time", *Nonlinear Analysis, Theory, Methods and Applications*, 2, pp. 355-373, 1978.
- [5] Carrero, M. and Pascali, E., "Il Problema del Rimbalzo Unidimensionale e sue Approssimazioni con Penalizzazioni non Convesse" *Rend. Mat.*, (6) 13, 4, pp. 541-553, 1980.
- [6] Lötstedt, P., "Mechanical Systems of Rigid Bodies Subject to Unilateral Constraints", *SIAM Journal of Applied Mathematics*, Vol. 42, No. 2, pp. 281-296, 1982.
- [7] Amerio, L. and Prouse, G., "Study of the Motion of a String Vibrating Against an Obstacle", *Rend. Mat.*, (2) 8, Ser. VI, pp. 563-585, 1975.
- [8] Amerio, L., "Su un Problema di Vincoli Unilaterali per l'Equazione non Omogenea della Corda Vibrante", *IAC (Inst. per le Applicazioni del Calcolo "Mauro Picone") Pubblicazioni*, Ser. D, 109, pp. 3-11, 1976.
- [9] Oden, J. T. and Martins, J. A. C., "Models and Computational Models for Dynamic Friction Phenomena", *Computer Methods in Applied Mechanics and Engineering*, 52, pp. 527-634, 1985.
- [10] Martins, J. A. C. and Oden, J. T., "Existence and Uniqueness Results for Dynamic Contact Problems with Nonlinear Normal and Friction Interface Laws", *Journal of Nonlinear Analysis* (to appear.)
- [11] Rabier, P. J., Martins, J.A.C., Oden, J.T., and Campos, L., "Existence and Local Uniqueness of Solutions to Contact Problems in Elasticity with Nonlinear Friction Laws", *International Journal of Engineering Sciences*, Vol. 24, No. 11, pp. 1755-1768, 1986.
- [12] Kikuchi, N. and Oden, J.T., *Contact Problems in Elasticity*, *SIAM Publications*, Philadelphia, 1987.
- [13] Back, N., Burdekin, M. and Cowley, A., "Review of the Research on Fixed and Sliding Joints", *Proc. 13th International Machine Tool Design and Research Conference*, ed. by S.A. Tobias and F. Koenigsberger, MacMillan, London, pp. 87-97, 1973.

- [14] Back, N., Burdekin, M. and Cowley, A., "Analysis of Machine Tool Joints by the Finite Element Method", Proc. 14th International Machine Tool Design and Research Conference, ed. by S. A. Tobias and F. Koenigsberger, MacMillan, London, pp. 529-537, 1974.
- [15] Lions, J. L., *Quelques Methods de Resolution des Prolemes aux Limites Non Lineaires*, Dunod, Paris, 1969.
- [16] Luenberger, D. G., *Optimization by Vector Space Methods*, Wiley, New York, 1969.

IN MATHEMATICS, No. 79, p.613-635 (Proc. Interdiscipline Symp. on FREE BOUNDARY PROBLEM: THEORY AND APPICATION, Vol.2). Fassans, A. and Primicerio, M. ed., Pitman, London.

- [12] Cheng, H.S. (1970), A Numerical Solution of the Elastohydrodynamic Film Thickness in an Elliptical Contact. *J.Lubri.Tech.*, 92,1,155-162.
- [13] Chipot, M. and Lustin, M. (1984), Existence and Uniqueness of Solutions to Reynolds' Equation. *Math.Rept.* 84-146, University of Minnesota, Minneapolis, Minnesota.
- [14] Ciarlet, P.G. (1978), THE FINITE ELEMENT METHOD FOR ELLIPTIC PROBLEMS. North-Holland Publishing Co., Amsterdam.
- [15] Cimatti, G. (1977), On a Problem of the Theory of Lubrication Governed by a Variational Inequality. *Appl.Math. & Opt.*, 3,2/3,227-242.
- [16] Cimatti, G. (1979), On the Existence of Solution for the Fundamental Problem of the Hydrodynamic Lubrication. *Int.J.Engng.Sci.*, 17,6,701-706.
- [17] Cimatti, G. (1980), A Free Boundary Problem in the Theory of Lubrication. *Int.J.Engng.Sci.*, 18,5,703-711.
- [18] Cimatti, G. (1984), On Certain Nonlinear Problems Arising in the Theory of Lubrication. *Appl.Math. & Opt.*, 11,3,227-245.
- [19] Coleman, C.J. (1980), A Finite Element Routine for Analysis Non-Newtonian Fluids, Part 1: Basic Method and Preliminary Results. *J.Non-Newtonian Fl.Mech.*, 7,4,289-301.
- [20] Coleman, C.J. (1981), A Finite Element Routine for Analysis Non-Newtonian Fluids, Part 2: The Extrusion of a Maxwell Fluid. *J.Non-Newtonian Fl.Mech.*, 8,3/4,261-270.
- [21] Curnier, A.R. and Taylor, R.L. (1982), A Thermomechanical Formulation and Solution of Lubricated Contacts between Deformable Solids. *J.Lubri.Tech.*, 104,1,109-117.
- [22] Day, C.P. (1974), Transient Elastohydrodynamic Lubrication by Finite Element Methods. M.S. Thesis, Cornell Univ., Jan., 1974.
- [23] Demkowicz, L., Devloo, P. and Oden, J.T. (1985), On an h-Type Mesh Refinement Strategy Based on Minimization of Interpolation Errors, *Comput.Meths.Appl.Mech.Engng.*, 53,1,67-90.
- [24] Demkowicz, L., Oden, J.T. and Strouboulis, T. (1984), Adaptive Finite Element Methods for Flow Problems with Moving Boundaries. Part1:

Variational Principles and A-Posteriori Estimates, *Comput Meths Appl. Mech. Engng.*, 46,2,217-251.

- [25] Derner et al, (1972), The Hollow Ended Roller-A Solution for Improving Fatigue Life in Assymmetrically Loaded Cylindrical Roller Bearing. *J.Lubri.Tech.*, 94,2,153-164.
- [26] Dorinson, A. and Ludema, K.C. (1985), **MECHANICS AND CHEMISTRY IN LUBRICATION**. Elsevier, New York.
- [27] Dowson, D. (1979), **HISTORY OF TRIBOLOGY**. Longman Group Ltd., New York.
- [28] Dowson, D. and Higginson, G.R. (1977), **ELASTO-HYDRODYNAMIC LUBRICATION**. Pergaman Press.
- [29] Dowson et al, (1979), **ELASTOHYDRODYNAMIC AND RELATED TOPICS. FIFTH LEEDS-LYON SYMPOSIUM ON TRIBOLOGY**, Mech. Engng. Publications, Bury St., Edmunds, Suffolk 9.
- [30] Duvaut, G. and Lions, J.L. (1972), **LES INÉQUATIONS EN MÉCANIQUE ET EN PHYSIQUE**. Dunod, Paris.
- [31] Eidelberg, B.E. (1974), Finite Element Analysis of Lubrication in Natural Joints. Ph.D. Thesis., Cornell University.
- [32] Eidelberg, B.E. and Booker, J.F. (1976), Application of a Finite Element Methods to Lubrication: Squeeze Films between Porous Surfaces. *J.Lubri.Tech.*, 98,1,175-179.
- [33] Friedman, A. (1982), **VARIATIONAL PRINCIPLES AND FREE-BOUNDARY PROBLEMS**. John Willey and Sons, New York.
- [34] Fujino, T. (1969), Analysis of Hydrodynamic and Plate Structure Problems by the Finite Element Method. JAPAN-U.S. SEMINAR ON MATRIX METH. STRUC. ANA. AND DESIGN, Tokyo, Japan, Aug.25-30, 1969, paper No. J5-4, p.725.
- [35] Glowinski, R., Lions, J.L. and Tremolieres (1976), **ANALYSES NUMERIQUE DES INÉQUATIONS VARIATIONNELLES**. Dunod, Paris.
- [36] Glowinski, R. and Marroco, A. (1975), Sur L'Approximation par Elements Finis D'Ordre Un, Et la Resolution, par Penalization-Dualite d'une Classe de Problems de Derechlet Non-Lineares. In *Revue Francaise d'Automatique et Recherche Operationnell*, R-2, 1975, p.41-76.
- [37] Grubin, A.N. and Vinogradova, I.E. (1949), Fundamentals of the Hydrodynamic Theory of Lubrication of Heavily Loaded Cylindrical

Surfaces. In INVESTIGATION OF THE CONTACT OF MACHINE COMPONENTS, Ketova, Kh.F. ed, Translation of Russian Book No.30, Central Scientific Research Institute for Technology and Mechanical Engineering, Moscow. Translation No.337, Dept. of Scientific and Ind. Res., London.

- [38] Gu, H.R. (1981), Finite Element Solution of the Elastohydrodynamic Lubrication Problem of Two Cylinders. M.S. Thesis (in Chinese), Xian Jiaotong Univ., Xian, China.
- [39] Guerra, F.M. and Becker, E.B. (1978), Finite Element Analysis of the Adaptive Method of Rezoning. *TICOM Report* 78-13, The University of Texas at Austin.
- [40] Hageman, L.A. and Young, D. (1981), **APPLIED ITERATIVE METHODS**. Academic Press., New York.
- [41] Hamrock, B.J. and Dowson, D. (1976-77), Isothermal Elastohydrodynamic Lubrication for Point Contact,
part 1, Theoretical Formulation, *J.Lubri.Tech.*, 98,2,223-229;
part 2, Ellipticity Parameter Results, *J.Lubri.Tech.*, 98,3,375-383;
part 3, Fully Flooded Results, *J.Lubri.Tech.*, 99,2,264-276;
part 4, Starvation Results, *J.Lubri.Tech.*, 99,1,15-23.
- [42] Hamrock, B.J. and Dowson, D. (1981), **BALL BEARING LUBRICATION, THE ELASTOHYDRODYNAMICS OF ELLIPTICAL CONTACTS**. John Wiley & Sons, Inc., New York.
- [43] Hayashi, H. and Taylor, C.M. (1980), A Determination of Cavitation Interfaces in Fluid Film Bearings Using Finite Element Analysis. *J.Mech.Engng.Sci.*, 22,6,277-286.
- [44] Huebner, K.H. (1971), Finite Element Analysis of Continuum Problems and Its Application to the Incompressible Lubrication Problem. *General Motors Research Publication GMR-1074*.
- [45] Huebner, K.H. (1974,a), Application of Finite Element Methods to Thermohydrodynamic Lubrication. *Int.J.Num.Meth.Engng.*, 8,1,139-165.
- [46] Huebner, K.H. (1974,b), Finite Element Analysis of Fluid Film Lubrication-A Survey. In **FINITE ELEMENTS IN FLUIDS**, Vol.2, p.225-254 (Int. Symp. FEM in Flow Problems, Jan.7-11,1974, Univ. College of Wales, Swansea). Galagher, R.H. et al ed., John Wiley & Sons, Great British.
- [47] Jian, B.N. (1986), Least Squares Finite Element Methods with Element-by-Element Solution Including Adaptive Refinement. Ph.D. Dissertation, The University of Texas at Austin.

- [48] Labouff, G.A. and Booker, J.F. (1985), Dynamically Loaded Journal Bearing: A Finite Element Treatment for Rigid and Elastic Surfaces. *J.Trib.*, 107,4,505-515.
- [49] Li, Z.C. and Dai, E. (1980), Numerical Methods for Calculating Pressure Distribution in Gas Bearing-Nonlinear Finite Element Method and Its Error Analysis for Solving Reynolds Equation (in Chinese). *ACTA Mechanica Sinica*, 1980, No.2, p.158-168.
- [50] Love, A.E.H. (1944), **A TREATISE ON MATHEMATICAL THEORY OF ELASTICITY**. 4th edit., Dover.
- [51] Mullen, R.L., Braun, M.J. and Hendricks, R.C. (1984), Finite Element Formulations of a Reynolds Type Equation for a Power Law Fluid. Presented on the Fifth Int. Symp. on FEM in Flow Problems, Jan.23-26,1984, Austin, Texas.
- [52] Oden, J.T. (1972), **FINITE ELEMENTS OF NONLINEAR CONTINUA**. McGraw-Hill, New York.
- [53] Oden, J.T. (1973), Approximation and Numerical Analysis of Finite Deformations of Elastic Solids. In **NONLINEAR ELASTICITY**, p.175-228. Dickey, D.W. ed., Academic Press, New York.
- [54] Oden, J.T. (1979), **APPLIED FUNCTIONAL ANALYSIS**. Prentice-Hall, Inc., Englewood Cliffs.
- [55] Oden, J.T. (1985), **QUALITATIVE METHODS IN NONLINEAR MECHANICS**. Prentice-Hall, Inc., Englewood Cliffs.
- [56] Oden, J.T. and Carey, G.F. (1982), **FINITE ELEMENTS**. Vol.4, Prentice-Hall, Inc., Englewood Cliffs.
- [57] Oden, J.T. and Kikuchi, N. (1980), Theory of Variational Inequalities with Applications to Problems of Flow through Porous Media. *Int.J.Engng.Sci.*, 18,10-A,1173-1284.
- [58] Oden, J.T. and Reddy, C.T. (1977), Finite Element Approximaion of a Two-Point Boundary Value Problem in Nonlinear Elasticity. *J.Elast.*, 7,243-263.
- [59] Oden, J.T. and Reddy, C.T. (1979), Finite Element Approximations of a Class of Highly Nonlinear Boundary-Value Problems in Finite Elasticity. *Numer.Funct.Anal. & Opt.*, 1,1,1-55.
- [60] Oden, J.T. and Reddy, J.N. (1976), **AN INTRODUCTION TO THE MATHEMATICAL THEORY OF FINITE ELEMENTS**. Wiley-Interscience, New York.

- [61] Oden, J.T., Strouboulis, T. and Devloo, P. (1986), Adaptive Finite Element Methods for the Analysis of Inviscid Compressible Flow: 1. Fast Refinement/Unrefinement and Moving Mesh Methods for Unstructure Meshes. *TICOM Report* 86-3, The University of Texas at Austin.
- [62] Oden, J.T. and Wu, S.R. (1985), Existence of Solutions to the Reynolds Equation of Elastohydrodynamic Lubrication. *Int.J.Engng.Sci.*, 23,2,207-215.
- [63] Oden, J.T. and Wu, S.R. (1986), An A-Priori Error Estimate of Finite Element Solutions for Elastohydrodynamic Lubrication. To be presented on The First World Congress on Computational Mechanics, Sept. 22-25, 1986, at the University of Texas at Austin.
- [64] Oh, K.P. (1974), The Application of a Gradient Method in the Elastohydrodynamic Problems. *ASLE Trans.*, 17,2,111-116.
- [65] Oh, K.P. (1984), Numerical Solution of Dynamically Loaded Elastohydrodynamic Contact as a Nonlinear Complementarity Problem. *J.Trib.*, 106,1,88-95.
- [66] Oh, K.P. and Goenka, P.K. (1984), The Elastohydrodynamic Solution of Journal Bearings under Dynamic Loading. *ASME paper* 84-Trib-20, presented at ASME/ASLE Lubrication Conference, Oct. 22-24, 1984, San Diego, California.
- [67] Oh, K.P. and Huebner, K.H. (1973), Solution of the Elastohydrodynamic Finite Journal Bearing Problem. *J.Lubri.Tech.*, 95,3,342-352.
- [68] Oh, K.P. and Rohde, S.M. (1977), Numerical Solution of the Point Contact Problem Using Finite Element Method. *Int.J.Num.Meth.Engng.*, 11,10,1507-1518.
- [69] Petrov, N.P. (1883), Friction in Machines and the Effect of the Lubricant. *Inzh.Zh.,St.Petersburg*, 1,71-140; 2,227-279; 3,377-436; 4,535-564.
- [70] Petrushevich, A.I. (1951), Fundamental Conclusions from the Contact-Hydrodynamic Theory of Lubrication. *Izv.Akad.Nauk. SSSR (OTN)*, 2,209.
- [71] Pinkus, M.K.E.O. and Sternlicht, B. (1961), **THEORY OF HYDRODYNAMIC LUBRICATION**. McGraw-Hill.
- [72] Pires, E.A.C.B. (1982), Analysis of Nonlocal Friction for Contact Problems in Elastostatics. Ph.D. Dissertation. University of Texas at Austin.
- [73] Reddi, M.M. (1969), Finite Element Solution of the Incompressible Lubrication Problem. *J.Lubri.Tech.*, 91,3,524-533.

- [74] Reddi, M.M. and Chu, T.Y. (1970), Finite Element Solution of the Steady State Compressible Lubrication Problem. *J.Lubri.Tech.*, 92,3,495-503.
- [75] Reynolds, O. (1886), On the Theory of Lubrication and Its Application to Mr. Beauchamp Tower's Experiments, Including an Experimental Determination of the Viscosity of Olive Oil. *Phyl.Trans.Roy.Soc.,A, London*, 177,157-234.
- [76] Rohde, S.M. and McAllister, G.T. (1975), A Variational Formulation for a Class of Free Boundary Problems Arising in Hydrodynamic Lubrication. *Int.J.Engng.Sci.*, 13,2,841-850.
- [77] Rohde, S.M. and Oh, K.P. (1975), A Unified Treatment of Thin Film Elastohydrodynamic Problems by Using Higher Order Element Methods. *Proc.Roy.Soc.London, A*, 343,315-331.
- [78] Shelly, P.D. and Ettles, C. (1973), A Finite Element Method for the Calculation of Locus Paths in Dynamically Loaded Bearings. *Proc.Instn.Mech.Eng.*, 187,5,79-86.
- [79] Showalter, R.E. (1979), **HILBERT SPACE METHODS IN PARTIAL DIFFERENTIAL EQUATIONS**. Pitman, London.
- [80] Shu, C.F. and Booker, J.F. (1983), Finite Element Analysis of Transient Elastohydrodynamic Lubrication. Presented at 10th Leeds-Lyon Symp. on Trib., Lyon, France, Sept., 1983.
- [81] Tansea, D.V. and Rao, I.C. (1966), Student Project Report on Lubrication, Royal Naval College, Dartmouth.
- [82] Tayal, S.P., Sinhasan, R. and Singh, D.V. (1981), Analysis of Hydrodynamic Journal Bearings Having Non-Newtonian Lubricants (Prandtl Model) by a Finite Element Method. *J.Mech.Eng.Sci.*, 23,2,63-68.
- [83] Tayal, S.P., Sinhasan, R. and Singh, D.V. (1982,a), Analysis of Hydrodynamic Journal Bearings Having Non-Newtonian Lubricants. *Trib.Int.*, 15,1,17-21.
- [84] Tayal, S.P., Sinhasan, R. and Singh, D.V. (1982,b), Analysis of Hydrodynamic Journal Bearings Having Non-Newtonian Lubricants Using the Finite Element Method. *ASLE Trans.*, 25,3,410-416.
- [85] Taylor, C. and O'Callaghan, J.F. (1972), A Numerical Solution of the Elastohydrodynamic Lubrication Problem Using Finite Elements. *J.Mech.Engng.Sci.*, 14,4,229-237.
- [86] Tieu, A.K. (1973), Oil-Film Temperature Distribution in an Infinite Wide Slider Bearing: An Application of the Finite Element Method.

J.Mech.Engng.Sci., 15,4,311-320.

- [87] Tieu, A.K. (1974), A Three-Dimensional Oil Film Temperature Distribution in Tilting Thrust Bearings. *J.Mech.Engng.Sci.*, 16,2,121-124.
- [88] Tieu, A.K. (1975), A Numerical Solution of Finite Width Thrust Bearings, Taking into Account Viscosity Variation with Temperature and Pressure. *J.Mech.Engng.Sci.*, 17,1,1-10.
- [89] Tower, B. (1883), First Report on Friction Experiments (Friction of Lubricated Bearings). *Proc.Inst.Mech.Eng.*, London, 632-659.
- [90] Wada, S. and Hayashi, H. (1971), Hydrodynamic Lubrication of Journal Bearings by Pseudo-Plastic Lubricants. Pt1, Theoretical Studies; Pt2, Experimental Studies. *Bull.JSME*, 14,268-286.
- [91] Wada, S., Hayshi, H. and Migita, M. (1971), Application of Finite Element Method to Hydrodynamic Lubrication Problems. Pt1, Infinite Width Bearings; Pt2, Finite Width Bearings. *Bull.JSME*, 14,1222-1244.
- [92] Willis, T. and Seth, B. (1977), An Application of the Finite Element Method to EHD Lubrication. *ASLE Trans.*, 20,4,340-346.
- [93] Wu, S.R. (1986,a), A Penalty Formulation and Numerical Approximation of the Reynolds-Hertz Problem of Elastohydrodynamic Lubrication. *Int.J.Engng.Sci.*, 24,6,1001-1013.
- [94] Wu, S.R. (1986,b), An Application of Penalty Method in Finite Element Solutions to Elastohydrodynamic Lubrication Problems. Presented at the 10th U.S. National Congress of Applied Mechanics, June 16-20, 1986, University of Texas at Austin.
- [95] Young, D. (1971), ITERATIVE SOLUTION OF LARGE LINEAR SYSTEMS. Academic Press, New York.
- [96] Zuk, J. (1976), Compressible Seal Flow Analysis Using the Finite Element Method with Galerkin Solution Technique. *ASLE. Trans.*, 19,1,61-71.